Association between mode of birth, staffing and structural characteristics in NHS trusts with maternity services in England (2010/11)

Gerova, Vania Nikolova

Awarding institution:
King's College London

Association between mode of birth, staffing and structural characteristics in NHS trusts with maternity services in England (2010/11)

Vania Gerova

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

Florence Nightingale School of Nursing and Midwifery

King's College London

May 2014

ABSTRACT

Background

Growing international research evidence, mainly from the acute general service sector, suggested that there was a strong link between nurse staffing and patient outcomes. There was a gap in the literature addressing other clinical and non-clinical workforce groups outside acute hospitals.

Aim

To investigate the relationship between mode of birth and maternity staffing levels in NHS trusts in England, after accounting for maternal socio-demographic characteristics, individual clinical risk and structural characteristics including type and configuration of trusts.

Method

This cross sectional study used Hospital Episode Statistics (HES) 2010/11 and NHS Information Centre 2010/11 maternity workforce datasets. The study population comprised women aged 15-45, who were nulliparous and had a term, singleton, live birth (n=261,481 deliveries in 143 NHS trusts for emergency caesarean section and instrumental deliveries; and n=214,920 deliveries in 127 NHS trusts for normal birth). Multilevel logistic models were fitted separately for each outcome. Risk-adjustment for case mix included maternal age, ethnicity, IMD, gestational age, birth weight and NICE 2007 derived definition of clinical risk. Standardized FTE/birth ratios for obstetricians, midwives, healthcare assistants and other trust characteristics were used as trust level predictors. The percentages of the total variation in outcomes attributable to between trusts variation were calculated.

Results

For this sample of women only around 2% of the residual variation in outcomes was due to unobserved trust characteristics. Between trusts and for all women, the standardized consultant FTE/birth ratio was positively related to the probability of instrumental delivery (OR=1.08, 95%CI 1.03-1.13, $p<.05$), and the standardized midwives FTE/birth ratio was positively related to the probability of normal birth (OR=1.06, 95%CI 1.01-1.11, $p<.05$). 1 SD increase in FTE doctors increased the odds of emergency CS for high risk women by 5.1% (OR=1.05, 95%CI 1.01-1.10, $p<.05$); while 1 SD increase in FTE midwives increased the odds of normal birth for low risk women by 7.6% (OR=1.08, 95%CI 1.02-1.14, $p<.05$).

Conclusion

The analyses established significant independent effects of staffing on the three outcomes, although only a small percentage of the total variability in the outcomes was attributable to variations between trusts. The positive association between midwifery staffing and normal birth has policy implications in terms of current and future investment in the profession. More than anything else, women's outcomes were determined by their characteristics and clinical risk. Other unaccounted for factors such as obesity, smoking, organisational culture and models of care may be able to explain further the variations in outcomes.

ACKNOWLEDGEMENT

I would like to thank my supervisors Professor Debra Bick and Professor Graham Cookson for their intellectual guidance in my academic work. Their encouragement, generous time, patience and support were invaluable for the completion of my doctoral work.

This academic work was funded by the National Nursing Research Unit, King's College London. I am very grateful to Professor Peter Griffiths and Professor Jill Maben for their trust.

Sincere thanks to Rod Gibson and Miranda Dodwell for their help with the data.

I am grateful to Trevor Murrells not only for his clear explanation of complex statistical concepts but for his great sense of humour and for reminding me what else matters in this life.

To my midwife Anathea Brooks from Elizabeth Garrett Anderson Hospital, UCH London, for giving a human face to this mainly desk-top work.

Warm thanks go to my high-spirited friend Evgenia Markova for all her help.

Special thanks to Helen, Naomi, Lynne, Lisa, Chloe and all the other mums for organising play time for my child and freeing some valuable time.

I am deeply grateful to my family, Gruda, Nikola and Georgi Gerovi and Alison and Robert Wynn for their support and love. Many thanks go to Mark and Kate Wynn and to Rob and Sarah Wynn.

Finally, I would like to thank my husband Gerard Wynn for his patience, support and constant supply of positive energy. This thesis is dedicated to my daughter Nicole Wynn for her endless inspiration and playful distractions.

ABSTRACT **2**

ACKNOWLEDGEMENT **3**

LIST OF TABLE **11**

LIST OF FIGURES **14**

1 CHAPTER 1 INTRODUCTION **15**

2 CHAPTER 2 MATERNITY WORKFORCE **28**

2.1 A RECENT HISTORY OF CHILDBIRTH IN THE UK 29

2.2 SAFETY, QUALITY AND EFFICIENCY OF MATERNITY SERVICES 35

2.2.1 High quality, safe, effective care 35

2.2.2 Minimising costs 39

2.3 POLICY SUPPORTING NORMAL BIRTH 44

2.4 CONTINUITY OF CARE 47

2.4.1 Woman-centred care 49

2.4.2 Midwife-led continuity of care 53

2.4.3 One-to-one care 56

2.5 ORGANISATION OF MATERNITY SERVICES 58

2.6 DEMAND-SIDE PRESSURES ON MATERNITY WORKFORCE 61

2.6.1 Birth trends and mothers' socio-demographic characteristics 61

2.6.2 Rise in complex pregnancies 63

2.7 SUPPLY-SIDE FACTORS AND STAFFING LEVELS 66

2.7.1 Midwives 66

2.7.2 Junior Obstetric Staff 77

2.7.3 Obstetricians 81

2.7.4 Matternity Support Workers 83

2.8 CONCLUSIONS 84

3 CHAPTER 3 MATERNITY CARE OUTCOMES 86

3.1 OBJECTIVES OF THE CHAPTER 86

3.2 BACKGROUND 87

3.3 NON-MATERNITY STAFFING LITERATURE 88

3.4 MEASURING QUALITY OF MATERNITY CARE 93

3.5 INDICATORS VS OUTCOMES 94

3.5.1 Structure/process/outcome 95

3.5.2 Desirable characteristics of indicators 97

3.6 STRUCTURE/OUTCOMES STUDIES 99

3.6.1 Staffing 99

3.6.2 Place of birth studies 105

3.7 PROCESS/OUTCOMES STUDIES 109

3.7.1 Continuity of care, one-to-one midwifery care, models of care 110

3.7.2 Interspecialty differences, philosophy 114

3.8 OUTCOMES STUDIES – MORTALITY, MORBIDITY, OPTIMALITY 115

3.8.1 Maternal Mortality *116*

3.8.2 Maternal Morbidity *117*

3.8.3 Optimality *124*

3.9 PROCESS/OUTCOME INDICATORS SELECTED FOR THIS THESIS 126

3.9.1 Caesarean section *126*

3.9.2 Instrumental delivery *135*

3.9.3 Normal birth *140*

3.10 CONCLUSIONS 145

4 CHAPTER 4 METHODS 147

4.1 OBJECTIVES OF THE CHAPTER 147

4.2 DATA SOURCES OVERVIEW 147

4.3 THE ENGLISH HOSPITAL EPISODE STATISTICS (HES) DATASET 151

4.3.1 Maternity tail *154*

4.3.2 Quality of HES and maternity tail data *154*

4.4 CLEANING OF HES 163

4.4.1 Multiple records *163*

4.4.2 Gestational age *163*

4.4.3 Parity *164*

4.4.4 Multiple births *164*

4.4.5 Maternal age *165*

4.4.6 Woman's ethnicity *165*

4.4.7 Birth weight *166*

4.5 NHS IC WORKFORCE STATISTICS, ENGLAND: 2010-11 166

4.6 DATA ON TRUSTS CHARACTERISTICS 171

4.6.1 Merging of trusts *173*

4.7 POPULATION USED IN THIS THESIS 173

4.8 SELECTED INDICATORS/OUTCOMES 176

4.8.1 Birth mode *176*

4.8.2 Normal birth *178*

4.8.3 Emergency Caesarean section *179*

4.8.4 Instrumental delivery *179*

4.9 BACKGROUND VARIABLES 180

4.9.1 Maternal age *180*

4.9.2 Woman's ethnicity *180*

4.9.3 Index of Multiple Deprivation (IMD) *182*

4.9.4 Birth weight *183*

4.9.5 Gestational age *184*

4.9.6 Smoking *184*

4.9.7 Definition of NICE (2007) clinical risk composite measure *185*

4.10 RISK ADJUSTMENT 188

4.11 STATISTICAL ANALYSES 191

4.11.1 The concept of multilevel modelling 191

4.11.2 Ordinary regression vs multilevel modelling 194

4.11.3 Basic multilevel model for continuous response variable 197

4.11.4 Modelling categorical outcomes 199

4.11.5 Single level logit 199

4.11.6 Multilevel logit model – Generalized Linear Mixed Model 201

4.11.7 Latent presentation model and variance partition coefficient (VPC) 206

4.11.8 Estimation procedures 207

4.11.9 Overdispersion and underdispersion 209

4.12 THE MODELS 210

4.12.1 Modelling stages 211

4.12.2 Software 213

5 CHAPTER 5 RESULTS 214

5.1 RESULTS 214

5.1.1 Descriptive statistics HES 2010/11 214

5.1.2 Staffing FTE/birth ratios 223

5.1.3 Emergency caesarean section – multilevel model 226

5.1.4 Instrumental delivery - multilevel model 238

5.1.5 Normal birth – multilevel model 249

5.2 CONCLUSIONS 260

6 CHAPTER 6 MODELS EXTENSION 262

6.1 MISSING DATA ANALYSES 262

6.1.1 Emergency CS – missing data analysis 262

6.1.2 Instrumental delivery – missing data analysis 263

6.1.3 Normal birth – missing data analysis 263

6.1.4 Tables from missing data sensitivity analyses 264

6.2 CROSS-LEVEL INTERACTIONS 273

6.2.1 Emergency CS – cross-level interaction analysis 274

6.2.2 Instrumental delivery – cross-level interaction analysis 275

6.2.3 Normal birth – cross-level interaction analysis 275

6.2.4 CONCLUSIONS 275

6.2.5 Tables from cross-level interaction analyses 277

7 CHAPTER 7 DISCUSSION 286

7.1 INTRODUCTION 286

7.2 PRINCIPAL FINDINGS 287

7.3 DISCUSSION 288

7.3.1 Selection of indicators 288

7.3.2 Selection of sample 289

7.3.3 Staffing and Trusts Characteristics and Outcomes 292

7.3.4 Women's characteristics and clinical risk 300

7.3.5 Variations in outcomes 304

7.4 CONTRIBUTION TO KNOWLEDGE 306

7.5 STRENGTHS AND WEAKNESSES OF THE STUDY 307

7.5.1 Generalisability and representativeness of the study sample 307

7.5.2 Variations in outcomes 308

7.5.3 Limitations of the data 310

7.6 POLICY AND CLINICAL IMPLICATIONS 313

7.7 FUTURE WORK 315

REFERENCE 317

APPENDIX I 344

APPENDIX II 345

APPENDIX III 347

APPENDIX IV 351

APPENDIX V 353

APPENDIX VI 359

LIST OF TABLE

Table 2.2-1 Number of maternal deaths by mode of delivery at 24 or more completed weeks of gestation; UK: 2006-08..... 44

Table 2.6-1: Live births in England, 1972-2012 61

Table 4.3-1 Number of valid records in HES by maternity key fields, 2010-11 and 2009-10..... 162

Table 4.5-1 HC & FTE for Selected Staff Groups in Maternity Services, England..... 168

Table 4.5-2 NHS Hospital and Community Health Services: Qualified Nursing, Midwifery & Support to Doctors and Nursing Staff in Maternity Services..... 169

Table 4.5-3 Medical Staff in O&G by Grade, FTE and Head Count 170

Table 4.6-1: Characteristics of NHS trusts with maternity services..... 172

Table 4.6-2 Grouped Trusts/PCTs..... 173

Table 4.8-1 Delivery Method (DELMETH) from HES maternity tail..... 177

Table 5.1-1: Descriptive Statistics; HES 143 trusts; 657,480 cases; April 2010/March 2011 215

Table 5.1-2: Demographic and clinical risk factors by birth outcomes HES April 2010/March 2011 . 217

Table 5.1-3: Descriptive Statistics, HES sample of 261,895 women; aged 15-44 years; nulliparous; at term (>=37w) with singleton, live births in 143 trusts..... 219

Table 5.1-4: Demographic and clinical risk factors by birth outcomes for women aged 15-44 years, nulliparous, with singleton, live births at term (>=37 weeks) 221

Table 5.1-5: NHS IC Maternity Workforce Data, at trust level, September 2010 224

Table 5.1-6: Standardized staff FTE/birth ratios at trust level 225

Table 5.1-7: Null Model, Emergency CS..... 227

Table 5.1-8: Fixed Effects, Level-1 predictors, Emergency CS *229*

Table 5.1-9: Fixed Coefficients, Level-1 predictors, Emergency CS *231*

Table 5.1-10: Fixed Effects, Level-1 & Level-2 predictors, Emergency CS *235*

Table 5.1-11: Fixed Coefficients, Level-1 & Level-2 predictors, Emergency CS *236*

Table 5.1-12: Intercept Variances from different stages of modelling Emergency CS *238*

Table 5.1-13: Null Model, Instrumental Delivery *239*

Table 5.1-14: Fixed Effects, Level-1 predictors, Instrumental Delivery *241*

Table 5.1-15: Fixed Coefficients, Level-1 predictors, Instrumental Delivery *242*

Table 5.1-16: Fixed Effects, Level-1 & Level-2 predictors, Instrumental Delivery *245*

Table 5.1-17: Fixed Coefficients, Level-1 & Level-2, Instrumental Delivery *246*

Table 5.1-18: Intercept Variances from different stages of modelling Instrumental Delivery *248*

Table 5.1-19: Null Model, Normal Birth *249*

Table 5.1-20: Fixed Effects, Level-1 predictors, Normal Birth *252*

Table 5.1-21: Fixed Coefficients, Level-1 predictors, Normal Birth *253*

Table 5.1-22: Fixed Effects, Level-1 & Level-2 predictors, Normal Birth *256*

Table 5.1-23: Fixed Coefficients, Level-1 & Level-2 predictors, Normal Birth *257*

Table 5.1-24: Intercept Variances from different stages of modelling Normal Birth *259*

Table 6.1-1: Fixed Effects Level-1 & Level-2 predictors, Emergency CS, missing data model *264*

Table 6.1-2: Fixed Coefficients, Level-1 & Level-2 predictors, Emergency CS, missing data model *265*

Table 6.1-3: Random Effect (Level-1 & Level-2 predictors Emergency CS with missing data model). *266*

Table 6.1-4: Fixed Effects Level-1 & Level-2 predictors, Instrumental Delivery, missing data model. 267

Table 6.1-5: Fixed Coefficients, Level-1 & Level-2 predictors, Instrumental Delivery, missing data model..... 267

Table 6.1-6: Random Effect (Level-1 & Level-2 predictors Instrumental Delivery with missing data model) 269

Table 6.1-7: Fixed Effects Level-1 & Level-2 predictors, Normal Birth, missing data model 270

Table 6.1-8: Fixed Coefficients, Level-1 & Level-2 predictors, Normal Birth, missing data model 270

Table 6.1-9: Random Effect (Level-1 & Level-2 predictors) Normal Birth with missing data model) . 272

Table 6.2-1: Fixed Effects, Cross-Level Interactions in Emergency CS Model 277

Table 6.2-2: Fixed Coefficients, Cross-Level Interactions in Emergency CS Model 278

Table 6.2-3: Random Effect, Cross-Level Interactions in Emergency CS Model..... 279

Table 6.2-4: Fixed Effects, Cross-Level Interactions in Instrumental Delivery Model 280

Table 6.2-5: Fixed Coefficients, Cross-Level Interactions in Instrumental Delivery Model 280

Table 6.2-6: Random Effect, Cross-Level Interaction in Instrumental Delivery Model 282

Table 6.2-7: Fixed Effects, Cross-Level Interactions in Normal Birth Model 283

Table 6.2-8: Fixed Coefficients, Cross-Level Interactions in Normal Birth Model 283

Table 6.2-9: Random Effect, Cross-Level Interaction in Normal Birth Model 285

LIST OF FIGURES

Figure 1: Proportion of births by caesarean section 1989-2010..... 130

Figure 2: Variations in elective and emergency CS across NHS trusts in England..... 131

Figure 3: Trends in emergency and elective CS in England 2000/2011 132

Figure 4 Instrumental delivery trends, England 2000/11..... 140

Figure 5: Data Flow 175

Figure 6: Graphic presentation of GLMM, adopted from Kachman 2000:61..... 202

Figure 7: Caterpillar plot of estimated residuals in 143 trusts for Emergency CS (Null Model) 228

Figure 8: Caterpillar plot of estimated residuals in 132 trusts for Emergency CS, after risk adjustment (Level-1 predictors)..... 233

Figure 9: Caterpillar plot of estimated residuals in 143 trusts for Instrumental Delivery (Null Model) 240

Figure 10: Caterpillar plot of estimated residuals in 132 trusts for Instrumental Delivery, after risk adjustment (Level-1 predictors) 244

Figure 11: Caterpillar plot of estimated residuals in 127 trusts for Normal Birth (Null Model) 251

Figure 12: Caterpillar plot of estimated residuals in 121 trusts for Normal Birth, after risk adjustment (Level-1 predictors)..... 255

1 CHAPTER 1 INTRODUCTION

Growing international research evidence suggests that there is a strong link between nurse staffing and patient outcomes (Aiken et al. 2002; Needleman et al. 2002; Lankshear et al. 2005; Rafferty et al. 2007; Kane et al. 2007; Aiken et al. 2014; Commission for Healthcare Audit and Inspection 2005; Royal College of Nursing 2006). Research from the US acute care sector during the last decade has shown that higher patient-to-nurse ratios are related to higher risk-adjusted rates of mortality among surgical patients and failure to rescue (Aiken et al. 2002). Published research within acute hospitals in the UK (Rafferty et al. 2007) supported the US findings (Aiken et al. 2002), by showing a relationship between low nurse-to-patient ratios and higher mortality and complication rates, which was also associated with low staff satisfaction and well being.

Growth of research on nurse staffing levels and skill mix has been driven by global recognition of nurse shortages, demographic pressures, cost reductions and patients' safety concerns. The majority of research to date has been undertaken in the acute general service sector and considered the impact of 'poor' nurse staffing levels on patients, nurses and the service organisation (Flynn and McKeown 2009). The emphasis seemed to have been also on adverse patient outcomes in relation to nurse staffing levels and skill mix. These included in-hospital mortality, failure to rescue, pressure ulcers, medication errors, hospital acquired infections, post-operative complications, length of hospital stay (Lankshear et al. 2005; Rafferty et al. 2007; Kane et al. 2007; Tourangeau et al. 2007; Aiken et al. 2014). At the same time there was a gap in the literature addressing other clinical and non-clinical workforce groups outside acute hospitals.

Given the evidence from acute general medical care, it might be reasonable to expect that variations in maternity staff levels, skill mix, models of care and providers characteristics might also have an impact on the quality of maternity care. However, little empirical evidence existed on the relationship between maternity staffing levels,

maternity workforce and workplace characteristics and birth outcomes. Following is a brief review of the relevant literature.

Only a few academic studies have examined exactly the relationship between staffing and maternal outcomes. These included cross-sectional and descriptive studies.

One cross-sectional study was undertaken by Joyce et al. (2002), which investigated the association between maternity staff (consultant obstetrician and gynaecologist (O&G), junior obstetric staff and midwives numbers per 1000 deliveries a year) and birth outcomes (caesarean section rate (CS), instrumental vaginal delivery rate (IVD) and epidural for labour rate), based on 1994-96 data for all maternity units in the Thames Valley region. The study found that variations in adjusted CS rates were related to the levels of monitoring and the experience of the obstetric staff; CS rates were positively correlated with junior obstetric staff levels and not associated with midwifery staffing levels; and staffing levels appeared unrelated to either epidural or IVD rates. Variations in epidural and IVD rate between units were most significantly explained by socio-demographic factors. Another cross-sectional study (Gerova et al. 2010) used routinely collected maternity data, based on 615,042 women who gave birth in 144 English NHS trusts (2008/09), and found a lower probability of readmission within 28 days after discharge from the postnatal ward associated with particular factors. These factors were: higher numbers of full time equivalent (FTE) midwives per birth; higher ratio of consultant obstetrician FTE to midwives FTE; and a higher ratio of consultant midwives FTE to midwives. However a limited selection of predictors was available: age and ethnicity of mother; Carstairs deprivation index; Charlson co-morbidity index; delivery type; professional overseeing the birth; number of admissions in the previous 12 months; pre-and post-birth length of stay. Therefore the risk adjustment was limited[1].

[1] The study used *Admitted Patients HES* data from Dr Foster for the period April 2008 – March 2009. Variables not included in the risk model: previous delivery type, parity, multiple pregnancies, multiple births, gestational age and co-morbidities, as the *Admitted*

Regarding descriptive studies, Ashcroft et al. (2003) focused on the labour wards of seven maternity units in the North West of England, examining the relationship between 'latent failures' ('accidents waiting to happen') and midwifery staffing, deployment and training, using a prospective semi-structured observational design. Situations were observed where shortages and poor deployment of midwives resulted in risk situations for mother and baby. One of the most disturbing observations was that 'near misses' because of staff shortages occurred frequently (on average one in every 2.5-5 days, usually in units with most deliveries and complications) and stayed unreported. 'Near misses' were defined as events which under different circumstances could become an accident. Another retrospective, descriptive study from the United States (Hall et al. 2009) looked at associations between the nursing care process (hourly quantity of nursing care received by each patient, calculated from documented nursing observations and interventions), patient characteristics (age, race, marital status, number of previous births, history of previous CS, augmented labour, weeks gestation, ICD-9 admission diagnosis and peak acuity, represented by labour level severity) and patient outcomes (duration of three stages of labour; labour complications; CS delivery; fetal distress; patient length of stay and cost of care) at patient and unit level. The study postulated that provision of additional nurse resources at key stages of labour improved labour progression and outcomes for women and their infants.

Other UK and international academic studies and systematic reviews have examined maternity outcomes, but not in relation to maternity staffing levels, for example examining pre and postnatal units, place of birth, models of care, continuous one-to-one support during labour and perception of safety.

Patients HES version did not contain the maternity tail, where some of the clinical data was available nor checked for women obstetric history from previous years in HES.

Forster et al. (2006) examined the impact of staffing[2] on the quality of postnatal care in one state in Australia. Smith et al. (2009) looked at perceived safety in maternity services. Hatem et al. (2008); Sandall et al. (2013) and Devane et al. (2010) examined differences in outcomes for childbearing women and their infants such as morbidity and mortality, effectiveness and psychosocial outcomes between midwife-led and other models of care; and Devane et al. (2010) assessed the potential cost-effectiveness of midwife-led maternity care in the UK. Hodnett et al. (2013) established the benefits for women and their infants of continuous one-to-one support during childbirth. Hodnett et al. (2010) and *Birthplace* in England (2011) compared maternal and neonatal outcomes by place of birth. However, these studies were not able to provide information on the effect of changes in staffing levels on birth outcomes or to guide policy on staffing levels.

Non-academic reports have found shortcomings in the quality of maternity services. The Care Quality Commission[3] raised concerns in their review of maternity services in 2008 *Towards Better Births*. The report stated that in some trusts "*levels of staffing were well below average, indicating that they may have been inadequate*" (HCC 2008). The review found: wide variations in staffing levels between trusts; variations in clinical outcomes; poor attendance at in-service training courses; variations in provision of supervision to midwives; and evidence of cultural separation between doctors and midwives. Similarly, the King's Fund independent inquiry in 2008, *Safe Births: Everybody's Business*, found shortages of midwives; inadequate consultant cover; inexperienced doctors; and issues related to training, appropriate skills deployment and communication between maternity teams.

[2] The study identified significant issues related to inadequate staff/patient ratios; staffing mix; patient mix; prioritisation of birth suites over postnatal units; and the use of non-permanent staff. Staff/patient ratios and recruitment/retention of midwives in rural areas was a major issue.

[3] Formerly the Healthcare Commission

The studies reviewed so far were concerned with maternity workforce, models of care, place of birth and outcomes. A separate literature has examined relevant policies. Policy studies have picked up on the key objectives of quality, safety, patient-centeredness and satisfaction with care, and the underlying aim of achieving the best outcomes for mother and baby with the least intervention possible.

Regarding the policy objective of quality, in a seminal report, the US Institute of Medicine (2001) defined quality of care as safe, effective, patient-centred, timely, efficient and equitable. This definition was adopted in NHS England (not just maternity services) by Lord Darzi (2008).

Regarding safety in England, The King's Fund report *Safe Births: Everybody's Business* (O'Neill 2008) recognised that for the majority of women in England, childbirth was safe for both mother and infant but some births were not as safe as they should or could be.

Although only a small number of British women die in childbirth, numbers have not declined in recent years. For example, mortality did not decline in 2003-2005, compared with the previous, triennial enquiry, according to the Confidential Enquiry into Maternal and Child Health (CEMACH 2007); there was however a small decline in 2006-2008 of the overall maternal mortality and larger reductions in deaths from some clinical causes (CMACE 2011). Regarding morbidity, a study jointly published by the Royal Colleges responsible for regulating maternity care standards found an increase in adverse obstetric events (RCOG 2007, *Safer Childbirth*). In addition, statistics show major changes in the last 30 years in England with respect to operative births, which carry increased risks of poorer health outcomes for mother and infant (RCOG 2013). NHS maternity statistics show that around 36% of women had operative delivery in 2010-11, compared with 20.2% in 1985; the caesarean section rate increased from 9% in 1980 to 25% in 2011-12 (NHS IC 2012); instrumental delivery has remained relatively stable in the last 30 years at 10-13% with an increase in vacuum extraction (NHS IC 2012); and spontaneous vaginal birth

declined from 75.4% in 1985 to 61.8 in 2010-11[4]. The issue of safety has contributed to rising maternity litigations[5]. There was approximately one claim for every 600 births in 2012-13; while a fifth of all the spending on maternity services was for clinical negligence cover (appr. £700 per birth) (NAO 2013).

Several complex and inter-related contributing factors were acknowledged by maternity care professionals (80% midwives) in England as affecting safety in maternity (Smith et al. 2009). Amongst them were staffing shortages; inadequate skill-mix and training; the rise of women with complex medical and social needs; poor management and team working.

A separate literature in UK has specifically exposed variations in outcomes across providers of maternity services (Paranjothy et al. 2005; Bragg et al. 2010; RCOG 2013; RCM 2011; CMACE 2011; and National Audit Office 2013). These studies and reports confirmed variations in CS rates and other outcomes across NHS providers with maternity services in England/Wales, and highlighted the importance of case-mix adjustment as some of the observed variations may be due to differences in the demographic and clinical characteristics of women; adjusting for case mix therefore allowed for a valid comparison between providers. These studies hypothesized that staffing levels; organisational factors and culture; women's preferences; and clinician's attitudes may explain the observed variations in outcomes across providers after case mix adjustment. However, the studies did not formally model staffing levels in relation to outcomes.

[4] Source: NHS Maternity Statistics, the Health and Social Care Information Centre 2011. http://www.hscic.gov.uk/pubs/maternity1011, Table 8. Please note that methodology was changed in 2006/07, with method of delivery derived from the OPCS codes in English HES, thus data is not completely comparable with previous years.

[5] Approximately 20% of all claims and approximately 60% of all payments relate to claims arising out of birth (NHS Litigation Authority 2006). Maternity care accounted for a third of the clinical negligence bill in 2012-13 (NAO 2013).

Regarding patient-centeredness, in recent years, maternity policies in England have followed the principles of:

- a 'national choice guarantee' policy for all women depending on their circumstances, i.e. choice of place of birth (at home; in a midwife-led unit; in an obstetric unit); choice of type of antenatal and postnatal care; and choice of how to access maternity care (*Maternity Matters*, DoH 2007; NICE clinical guideline 55, 2007). To support the best possible care, the Government in 2008 announced extra funding for maternity, totalling £330 million over the following three years. At the same time moves to centralise birthing facilities in recent years with a rationale of cutting costs and improving safety and efficiency are reducing the choices for women and may have a negative impact on the outcomes of care and experiences of women (Devane et al. 2010);
- continuity of care with its potential to improve maternal and infant outcomes (Hatem et al. 2008);
- one-to-one midwifery support during labour (NICE clinical guideline 55, 2007).

A further literature has established the workforce challenges regarding achieving these priorities of quality, safety, woman-centeredness, continuity of and one-to-one care and choice. Both demand and supply-side challenges can be identified, as well as the impacts of the global financial crisis.

On the demand side, the changing profile of women was seen by clinicians as placing additional demands on care provision and clinical involvement (RCM 2005; RCOG 2007; CEMACH 2007; CMACE 2010; King's Fund 2011; and NAO 2013). These changes included: rising birth rates; care for older first-time mothers; care for women with more complex pregnancies and health needs such as obesity and diabetes; language barriers for a rising immigrant population; the impact of an increase in social inequalities on pregnancy and postnatal health. Regarding care for women with co-morbidities, these can increase intervention rates, with implications for

staffing levels. Length of stay in hospital is also seen as having implications on need for additional staffing levels (RCM 2005; King's Fund 2011).

Policies such as woman-centeredness have created an additional challenge. For example, offering choice over mode and place of birth can be more labour intensive. Maternity policies in England currently promote 'normal birth' (i.e. birth without medical interventions) as a desirable outcome for women (Department of Health 2007). Inadequate midwifery staffing levels were cited as an obstacle in achieving 'normal birth' (Page 2003); in providing one-to-one care (Reform 2005; NAO 2013) and to safe care in general (Smith et al. 2009). Concerns over low numbers and lack of experience among doctors (Smith et al. 2009) were also raised. The latest National Audit Report (2013) revealed that though three quarters of obstetric units have achieved 60 hours (per week) or more of consultant obstetrician presence on labour wards in 2012; 53%[6] of units were not compliant with the levels of consultant presence recommended by RCOG (2007) and endorsed by the NHS Litigation Authority. *Towards Better Births* (HCC 2008) review of maternity services in England reinforced the need for increasing the number of midwives. There was a promise by the Secretary of State for Health in 2008 to support the recruitment in the NHS of an extra 4000 (3400 full-time equivalent) midwives by 2012 (NHS WRT 2009). The RCM submitted a memorandum (HS41) to Parliament in 2011, establishing that the country needed the equivalent of 4700 extra FTE midwives (CfWI 2011). This estimate was based on *Birthrate Plus* workforce planning tool. The September 2013 figures showed that there were the equivalent of 1158 more full-time midwives in the NHS in England in that month, compared to September 2010 (HSCIC 2014). This suggests that the RCM message over the last few years on the need for investment in maternity services and generally in more midwives might have been partially acknowledged by the decision-makers.

[6] 100% of the largest units (with 5,000+ births per year) were non-compliant in providing the recommended 168 hours consultant presence per week.

The main supply-side pressures on the maternity workforce have been outlined in a series of reviews (*Midwifery 2020*, 2010; Reform 2005; HCC *Towards better births* 2008; NHS Workforce Review Team 2009/10; RCM 2005). Specifically, there is a shortage of midwives as a result of a number of factors, including: aging and retirement of midwifery workforce; fewer practicing than registered midwives; and levels of sickness/maternity leave. In addition, more midwives were opting to work part-time, meaning that although the headcount had increased in recent years, the full-time equivalent (FTE) of working midwives has not. Other impacts on the supply side relating to midwives are: recruitment, retention and staff well-being; changes in educational system (introduction of degree level for midwives); re-defining the roles and broadening the scope of the responsibilities of midwives from antenatal through intrapartum, postnatal and community care, with focus on normal birth (and the associated demand pressures of home births, one-to-one and continuity of care). Supply-side pressures impacted on the changing roles of maternity support workers (Sandall et al. 2007; Stout 2007; RCM 2010).

Supply-side impacts regarding obstetric medical staff, include: shortening the training and restructuring of the career paths for new doctors (*Modernising Medical Careers* 2005[7]), the reduction in working hours for junior doctors due to the European working hours regulations[8] (*Temple Report* 2010); and the call for 24-hour consultant presence[9] on labour wards to compensate for reduction in trainee numbers and hours and a rise in case mix complexities (*The Future Role of the Consultant* RCOG 2005; *The Future Workforce in Obstetrics and Gynaecology* RCOG 2009). Recognising the complexity of the issues affecting the maternity workforce, a House

[7] House of Commons Health Committee (2008) Third Report of Session 2007–08.

[8] European Working Time Directive (EWTD) was introduced in 1998 and fully implemented in UK in 2009 including for junior doctors. It limits work to 48 hours per week averaged over 6 months.

[9] *Safer Childbirth (2007)* recommended 40 to 60 (for units with more than 5000 births per annum) hours per week presence of obstetricians on labour wards.

of Commons Health Committee report (2007) called for an integrated and cross-professional approach to workforce planning, aiming at having the "right people, with the right skills, in the right place, at the right time" (NHS Workforce Review Team (WRT) 2009).

The global financial crisis has also posed its own challenges. NHS efficiency savings of £20 billion were planned by the end of 2013/14 (QIPP, DoH 2010) mainly through improved productivity, innovations, quality and prevention, which is a challenge to maternity services as well. The NHS is the largest employer in the UK and midwives and nurses represent the largest group of employees within the organisation. The current economic policies of public funding cuts and the focus on improving productivity in the public sector, including health, will have a profound effect on the NHS workforce, mainly because of the labour intensity of the health sector. The challenges to the NHS workforce of the abolition of mandatory retirement age are to be seen.

This policy literature, therefore, has identified pressing challenges including: shortages of midwives; public spending cuts; an ongoing focus on improving productivity and on quality and safety of maternity services; patient-centeredness; recognised variations in outcomes across NHS providers with maternity services; and the high litigation rate for obstetrics.

Meanwhile, the academic literature regarding the relationship between maternity staffing and outcomes is sparse.

This thesis aimed to investigate the relationship between maternity staffing levels (medical and non-medical), structural characteristics (type and configuration) of NHS trusts with maternity services and three selected birth modes – emergency CS, instrumental delivery and normal birth in NHS trusts in England in 2010/11. It aimed to answer four specific questions:

1. What is the relationship between emergency CS, instrumental delivery and normal birth and maternity staffing levels (FTE/birth ratios) in NHS trusts in England in 2010/11, after accounting for maternal socio-demographic characteristics, individual clinical risk and structural characteristics including type and configuration of trusts?
2. Is this relationship, if any, different across the three indicators?
3. To what extent do maternal socio-demographic characteristics and clinical risk factors determine the likelihood of emergency CS, instrumental delivery and normal birth?
4. To what extent are variations in emergency CS, instrumental delivery and normal birth between NHS trusts with maternity services in England explained by differences in staffing levels and trusts' characteristics?

To answer these questions a cross sectional study using routinely collected data was undertaken. Two major datasets were matched at trust level (English HES 2010/11 and Maternity Workforce Dataset from NHS IC 2010/11). Multilevel logistic models were fitted separately for each outcome: emergency CS; normal birth and instrumental delivery. This approach accounts for the clustering of the data, examines the relationship between staffing levels and outcomes and tests the hypothesis that trusts and workforce differences in the provision of maternity care contribute to the variations in the outcomes between trusts. The study findings will add to the currently limited evidence on the relationship between staffing and outcomes in maternity.

THESIS STRUCTURE

Chapter II presents the main policies, demographic and staffing issues affecting maternity staffing groups.

Chapter III focuses on the literature and empirical evidence related to quality of healthcare and birth outcomes and the justification for the selection of the final outcomes to be investigated.

Chapter IV examines the data sources and multilevel modelling approach.

Chapter V presents the results of the models for emergency CS, instrumental and normal birth.

Chapter VI presents the results from sensitivity analyses.

Chapter VII discusses the principal findings and the strength and the limitations of the empirical results; considers the policy implications of the study; recommendations for future work and conclusions.

THE APPROACH IN THIS THESIS

This thesis did not aim to establish indicators of overall quality applicable to maternity care but aimed to examine if any association between structure (staffing levels, structural characteristics) and process/outcomes (the interventionist process indicators of emergency caesarean section and instrumental delivery and the outcome indicator - normal birth) exist. Furthermore, it aimed to investigate whether maternity staffing levels and structural characteristics explain variations in the selected process/outcome indicators?

The search strategy to identify literature on relationships across structure/process/outcomes related to maternity care, did not follow a specifically written algorithm and entailed database searches in CINAHL; Medline; MIDIRS; PubMed; Cochrane Library; Web of Science and Google Scholar for the period

2000-2012. Keywords used in the search included ‘midwife’, ‘midwifery/obstetric outcomes’, ‘birth outcomes’, ‘adverse obstetric event’, ‘severe obstetric morbidity’, ‘maternity/midwifery staff’, ‘obstetric/midwifery staffing levels’, 'maternity/midwifery staffing mix', ‘maternity/obstetric staff’.

Many of the policy documents were found on the relevant organisations web pages, including documents published post 2012. Identification of many academic publications was done via the bibliographic pages of important academic papers in the process of writing the literature review and the methodology chapters.

2 CHAPTER 2 MATERNITY WORKFORCE

This Chapter examines the main policies and supply and demand pressures affecting maternity services and maternity workforce in the NHS in England. The Chapter presents information on current health policies and trends in England in relation to: maternity workforce; recent birth trends; mothers' socio-demographic characteristics; organization of maternity services and maternity workforce characteristics (workforce demographic changes and role changes); supply and demand pressures; and productivity challenge.

The Chapter is intended to set the scene and to map the complexity of policies, demographic, economic and organisational issues impacting on the maternity workforce. It is important to note that this Chapter does not present an in-depth critique of the literature and is more descriptive than analytical. The following Chapter 3 provides a critical analysis of the academic literature on outcomes and staffing.

The Chapter aimed to answer the following questions:

1. What are the key policy drivers affecting maternity services and maternity workforce in England?
2. What are the main demand and supply-side factors affecting maternity workforce?

The first question explored policy issues related to maternity workforce, such as:

- Provision of woman-and baby-centred care: women's choice of place of birth (at home; in a midwife-led unit; in an obstetric unit), midwife-led continuity of care, provision of one-to-one care;
- Normality in birth and reduction of unnecessary interventions;
- Safety, quality and efficiency of maternity services;

- Organisation and re-organisations of maternity services (centralization of maternity units and shift of care from acute to community setting);
- NHS budgetary cuts and focus on productivity and innovations.

The second question on demand and supply-side factors will examine:

- Demand-side pressures on the maternity workforce: changes in population demographics - increased birth rates; older first-time mothers; rise in complex pregnancies (i.e. maternal obesity, co-morbidities) and social inequalities; normal birth policy; reduction of training and working hours for junior doctors; 24 hours presence of consultant-obstetrician on the labour ward; models of care; midwifery staffing requirements in obstetric and midwife-led units (the *Birthrate Plus* midwifery planning tool), home births and community midwives; maternity support workers roles and regulation;
- Supply-side pressures affecting future maternity workforce planning: aging midwifery workforce and retirement; recruitment, retention, flexible working and staff wellbeing; education changes; changing midwifery practices.

2.1 A RECENT HISTORY OF CHILDBIRTH IN THE UK

Throughout the 20th century there has been a trend in the developed world for reduced maternal and perinatal risk of dying and improvement in the population's general health and nutrition. Maternal mortality was on the rise in England and Wales between 1900 and 1937 (to over 40 deaths per 10,000 births), while infant mortality had been gradually declining (from nearly 170 deaths per 1,000 live births in 1900 to under 35 deaths per 1,000 by 1948) (Davis 2013). After 1948, both rates continued to decline. Maternal deaths related to pregnancy, childbirth and the puerperium[10] in England and Wales were 1 in 15,000 in 2012 (ONS 2013a in NAO

[10] Within 6 weeks following birth.

2013). The perinatal mortality rate[11] was 7.5 per 1,000 births in England; 6.6 in Wales; 6.9 in Scotland and 6.4 in Northern Ireland (ONS 2013b in NAO 2013).

The evolution in the safety of childbirth was driven by the availability of antibiotics, blood transfusion and improved anaesthesia in operative procedures. Perinatal mortality came more into focus in 1960s with the introduction of new technologies, such as ultrasound, antenatal cardiotocography, amniocentesis and fetoscopy, although the reduction in perinatal mortality in the developed world was more related to socio-economic factors (educational, social, and health systems), changes in nutrition and trends in reproduction rather than advances in obstetric medicine (Chamberlain 1991).

With the foundation of NHS in 1948, there was a renewed interest in maternal health but in the following 26 years there was a lack of "*clear, universally agreed vision for maternity care"* (Davis 2013:3). During that time the provision of maternity care followed the 'tripartite system' of the NHS, divided between hospital services, General Practitioners (GPs) and maternal clinics run by the local authority health services. In recognition of the assumed technology-based fall in maternal mortality the preferences for giving birth gradually shifted towards consultant-led care in hospitals, where most of the women giving birth were still cared for by midwives. The hospitalisation of births was accelerated by two maternity services reviews - the *Cranbrook Report* in 1959[12] and the *Peel Report* in 1970[13] and in the 1970s 95% of

[11] Stillbirths and babies dying within 7 days of birth.

[12] Maternity Services Committee (1959) *the Cranbrook Report* recommended that 70% of all births should be provided for by maternity services in hospitals; while for the other 30% of women, after an appropriate consideration for giving birth at home or hospital, it was deemed safe to give birth at home (Davis 2013).

[13] The recommendations in this report were based on findings from the Reports of the Confidential Enquiry into Maternal Deaths and stated that hospital facilities should be available for all women to give birth in hospitals. Maternity and Midwifery Advisory Committee (1967) *the Peel Report* recommended 100% hospital deliveries with medical and midwifery care provided by teams of consultants, GPs and midwives (Davis 2013).

women were giving birth in hospitals (Davis 2013). The *Peel Report* inferred that hospital births were safest but was criticised for lack of evidence and for not considering women's views about their experiences and preferences (Davis 2013). This dramatic shift in hospitalisation of birth is still evident today. In 2012, 87% of women in England gave birth in a hospital obstetric unit (NAO 2013), albeit the configuration of maternity services has been changing, with local variations[14] in different types of birth settings. 11% of births in 2012 in England were in midwifery-led units (an increase of 7% from 2006-07), while 2.4% were home births in 2011 (a decline of 0.4% from 2007) (NAO 2013).

New technologies were introduced in the 1970s, for example: antenatal testing and monitoring (using ultrasound), which from diagnostic tools for high-risk pregnancies, became routinely used in the 1980s. The technological and medical advancements improved diagnoses in high-risk pregnancies but also contributed to a rise in caesarean section rates, oxytocin drugs to induce labour and routine use of episiotomy. This trend was confirmed by the National Birthday Trust Fund in a national survey of all births in the UK within one week in 1970 (*British Births*[15]), which established the increased use of these interventions compared to rates in a previous survey in 1958 (Davis 2013). The report was originally looking at links between socio-economic status and infant health (Elliott and Shepherd 2006). It acknowledged the reduction in perinatal mortality (from 33 per 1,000 in 1958 to 23 per 1,000 in 1970), mainly because of reduced deaths due to pneumonia, birth trauma, pulmonary haemorrhage, and haemolytic disease; while low birth weight (less than 2500g at birth) remained one of the main causes of perinatal mortality (Editorial, BMJ 1976). It also considered differences in perinatal mortality rate in the contexts of reduced numbers of elderly mothers (>35 years) but a rise in number of

[14] 21% of mothers in the East of England gave birth outside obstetric unit (midwifery-led or home births), while in East Midlands the rate was 4% (NAO 2013).

[15] British Births 1970, The First Week of Life. London, Heinemann, 1975 (in BMJ Editorial 1976).

mothers aged under 20 years; an increase in young immigrant mothers; and widening of social inequalities.

In the mid 70s, a public debate ensued about obstetric practices and particularly the controversial routine use of induction and acceleration of labour (Davis 2013). Cochrane (1979) in his review of the medical profession pointed out that new expensive innovations (such as induction of labour, ultra-sound, fetal monitoring, and placental function tests) were introduced by obstetricians during pregnancy, labour and birth without rigorous evaluations and resulting in higher costs per birth. He criticised the obstetric specialty for having the "*poorest record in evaluating their practices*" (cited in Enkin et al. 2006:265); for their delayed interest in randomised clinical trials[16]; for failing in the 1960s to randomise low risk pregnant women to home or hospital, thus having nearly all of them delivering in hospitals (Cochrane 1979 in Enkin et al. 2006).

Consequently, in the 1980s and 1990s the medicalisation of hospital births began to be questioned by women and health professionals, backed up by evidence-based[17] research (Chalmers et al. 1986; Chalmers et al. 1989 and Enkin et al. 1989). The publication of "*Effective care in pregnancy and childbirth"* in 1989 edited by an Englishman (Chalmers), a Canadian (Enkin) and a Belgian (Keirse) revolutionised the way maternity care was evaluated, but as Enkin et al. (2006) acknowledged, this seemed to have had little influence on obstetrical practice. The book was a result of a

[16] Randomised clinical trials were introduced to medicine in the 1950s. Randomised control trial is a study which tests a drug or a treatment by randomly assigning people to two (or more) groups: one (the experimental group) receiving the drug/treatment that is being tested, and the other (the comparison or control group) receiving an alternative drug/treatment, a placebo or no treatment. The differences in outcomes for the two groups are compared over time to establish the effectiveness of the experimental drug/treatment. When randomising, the groups should be similar in all aspects apart from the treatment they receive during the study.

[17] The term 'evidence-based' medicine was used for the first time in 1991 by Gordon Guyatt, of McMaster University in Hamilton, Ontario. It is the process of systematically finding, appraising, and using research findings which are used to justify clinical decisions.

thorough investigation of all the evidence for obstetric interventions and the roles of the professionals involved. It provided evidence from systematic reviews of randomized trials for useful practices and against those that were ineffective or harmful. The work was later computerised through the Cochrane Collaboration[18], which today is an indispensable global reference source for all maternity services.

Professional liability emerged as a new risk related to childbirth in the 1970s. The trend in obstetric litigation started in the 1970s and continued to rise in the 1980s in most developed countries and today is the biggest financial burden to the UK health system. Between 1995 and 2005, £1,513 million payments (including amounts set aside for unresolved cases) for obstetric related incidents were made by the NHS Litigation Authority (RCOG 2007). Fear of litigation encouraged a practice of defensive medicine – e.g. unnecessary tests and fetal monitoring were performed. The consequences of having a self-protective approach to providing care were that some treatments were modified to self-protect the physician instead of giving the most effective care based on women's needs, resulting in increased rather than decreased overall risk (Enkin 1994). The risk-management approach[19] to childbirth regarded birth as a life-threatening event which required monitoring and interventions; it proved to be beneficial in truly high risk situations but for a minority of women (Enkin 1994). Enkin wrote in 1994 "*Labor can now be started at will, monitored, augmented, and pre-empted*" (p.132) and questioned whether risk-management approach was appropriate for the majority of women for whom birth was a demanding but natural event. The view that pregnancy and childbirth were medical events requiring medical solutions became no longer acceptable.

[18] An international organisation in which, randomised controlled trials are appraised and reviewed by different professions. The Cochrane Database of Systematic Reviews contains regularly updated reviews on a variety of health and other issues and is available electronically as part of the Cochrane Library.

[19] Risks to the mother and baby (mortality, morbidity) and physician (litigations).

Midwifery-led non-interventionist care promoting 'normal' pregnancy and 'normal' birth was accepted as appropriate for uncomplicated (low risk) pregnancies. Midwives were recognised as "*experts in protecting, supporting, and enhancing the normal physiology of labor, delivery, and breastfeeding"* (Rooks 1999:370). Consultant-led interventionist care was perceived as better suited for women with recognised diseases (higher-risk pregnancies) and in case of complications. As medical training is focused on pathology[20] and obstetric medical care prepares for complications in pregnancy and birth (as any woman potentially could develop sudden complications), a clear distinction between pathological and risk factors[21] was needed to reduce unnecessary interventions. This required clarity of definition, interpretation and application of risk criteria in conjunction with well organised continuous risk assessment of women at the antenatal stage and during labour.

Midwifery and medical obstetrics are separate but complementary disciplines (Rooks 1999). Currently, medical and midwifery models of care are not mutually exclusive but complement each other. In the UK, midwives are not only recognised as the autonomous practitioners of normal labour and birth, but work also in teams with obstetricians, anaesthetists, paediatricians, nurses and support workers in care of women with complex pregnancies and labour; consultant obstetricians accepted the labour ward as part of their regular responsibilities as was recommended in the 1999 RCOG report *Towards Safer Childbirth* (in RCOG 2007).

Midwifery-led units for low-risk women (freestanding or alongside units) and obstetric-led maternity units, where midwives provide the majority of care for low and high-risk women, were introduced. A growing body of research found that births in midwife-led units were safe, effective and cost-effective (Hodnett 2002; Reinhaz et al. 2000; Rosser and Anderson 2000; Saunders et al. 2000; Sandall et al. 2013).

---

[20] Pathology – study and diagnosis of disease.

[21] Conditions, which are not pathological, but are associated with higher risk of complications.

Following this brief history is, first, a review of the key policy drivers affecting current maternity services, and second, a description of the main demand and supply-side factors affecting the maternity workforce.

2.2 SAFETY, QUALITY AND EFFICIENCY OF MATERNITY SERVICES

Safety, quality and efficiency have been major policy priorities for more than a decade, driven by concerns over minimising costs, including litigation, and consumer demands for a better service.

2.2.1 HIGH QUALITY, SAFE, EFFECTIVE CARE

Maternity care is the most common reason for hospital admission for women aged 15-59 years in the UK (HSCIC 2012). Provision of safe and high quality maternity care is therefore a core policy priority, as documented in government papers, academic articles and independent inquiries.

Regarding NHS in England, Lord Darzi's *Next Stage Review* (2008) adopted three of the quality dimensions (safe, effective, patient-centred) from US Institute of Medicine. The White Paper outlining strategy for the NHS in England *Equity and Excellence: Liberating the NHS* (DoH 2010a) affirmed the current coalition government commitment in relation to providing care that was safer, more effective, and which provided a better experience for patients. The Government outlined that the purpose of the NHS-funded care was improvement in quality and healthcare outcomes. A simplified definition of safety for maternity care was presented in *Safe Births: Everybody's Business* (King's Fund 2008), as: *"the avoidance, prevention and amelioration of adverse outcomes or injuries stemming from the process of health care"* (Vincent 2006).

The current Government has committed to focus on *"clinically credible and evidence-based outcomes"*, as the measure of quality of care across the NHS. The *NHS Outcomes Framework* is in the process of devising a comprehensive set of

indicators, based on quality standards including 150 indicators to be developed by NICE by 2015. The *NHS Outcomes Framework* was launched in April 2012, to focus on developing a set of national outcome goals, to drive the overall improvements in the NHS, and establish criteria against which to hold the NHS Commissioning Board to account. The framework was developed by the NHS Commissioning Board in consultation with clinicians, patients and the public.

Incentives for better quality would mean that payments to providers will be linked to outcomes, not just activity. The planned £20 billion savings by 2014 were to be reinvested to support improvements in quality and outcomes. The outcomes would replace process indicators and targets, which were seen as top-down, excessive bureaucratic measures.

The Government's aims, under the new approach of focusing on outcomes, included improving effectiveness, by reducing mortality, morbidity and emergency re-admissions, and through patient-reported outcome measures (PROMs). Another objective was to increase safety, such as lower rates of adverse events and avoidable deaths. The government also aimed to improve patient experience and outcomes such as reduced waiting times and fewer cancelled operations (DoH 2010a).

Regarding maternity service provision in England specifically, delivering a choice of safe, accessible, high quality care was a central objective, as expressed in the Department of Health's earlier document, *National Service Framework for Children, Young People and Maternity Services* (2004) and *Maternity Matters* (2007). Maternity services performance was seen as "*a touchstone of whether we are delivering quality based on patient safety, effectiveness of care and patient experience"* (David Nicholson NHS Chief Executive, in his NHS annual report 2008/09, DoH 2009).

Lord Darzi's *High Quality Care for All* (DoH 2008) identified nurses and midwives as the key professions to achieve the aim of delivering high quality care that is *"effective, safe and patient-centred".* The recognition that nurses and midwives

were at the heart of delivering the *"high quality, compassionate care that everyone wants"* was also a focus of the *Front Line Care* (2010)[22] report, whose recommendations included identifying better outcome indicators for midwifery. *Midwifery 2020* (2010) aimed to establish an informed vision of how midwives could contribute to achieving quality, cost-effective maternity services for women, babies and families.

Having identified midwives as the key workforce to provide high quality, safe maternity care, several hurdles were identified (RCM 2005; RCOG 2007; CEMACH 2007; CMACE 2010; King's Fund 2011; and NAO 2013). Specifically, the maternity services were expected to provide better outcomes including safe and positive experiences for mothers and babies, while coping with pressures including: increased demand; complex pregnancies; demographic changes; financial constraints; and the need for service reconfigurations affecting maternity workforce levels, staffing mix, roles and settings.

Maternity staffing requirements were informed by the joint publication from the Royal Colleges in 2007 *Safer Childbirth* (RCOG 2007), which aimed to establish the minimum standards for the organisation and delivery of care in labour. Standard 10 focused on childbirth outcomes (including normal births without interventions; emergency CS – incidence and indications and instrumental births – ventouse, rotational or non-rotational forceps); while Standard 4 focused on provision of safe staffing levels (midwives, consultant obstetricians, junior obstetric staffing, junior medical staff – obstetricians, anaesthetists and paediatricians) for each birth setting; other standards included leadership; multidisciplinary working; communication; core responsibilities, training and education.

The *Commissioning for Quality and Innovation* (CQUIN) payment framework was introduced in 2009 to enable commissioners of health services to provide financial

[22] Prime Minister's Commission on the Future of Nursing and Midwifery in England.

incentives in support of local quality improvement goals. Its aim was to reward implementation of NICE quality standards; improvements of patient experience and patient-reported outcomes; while poor quality care would be penalised (imposing fines for a list of 'never events', such as wrong site surgery, from October 2010). The CQUIN framework included the following Indicators for Quality Improvement (IQI) related to maternity: indicators for smoking cessation during pregnancy, prevalence of breastfeeding at 6-8 weeks and access to maternity services by 12 weeks + 6 days (NHS IC 2013a). Indicators of quality therefore should measure a balance of aspects of harm or adverse outcomes, care which promotes health and patient derived measures of their experience of care.

Other indicators of quality recommended by *Midwifery 2020* (*Delivering Expectations* 2010) included: increasing normal birth rate, continuity of midwife-led care, reducing perineal trauma and increasing skin-to-skin contact.

More recently, the RCOG proposed significant changes in the way maternity services were structured, to allow a life-course approach to women's health, in their report *High Quality Women's Health Care* (RCOG 2011). That meant improving both maternity outcomes, and managing women's health over an individual's lifetime. These significant changes would involve establishing clinical networks combining primary, community, secondary and tertiary services; developing universally adopted clinical standards to reduce variations in care; reducing the numbers of medically staffed units to ensure safe service and provision of 24/7 medical obstetric service; and reconfiguration of hospitals to accommodate more midwife-led care so that timely and safe care was possible from the multidisciplinary teams.

The latest *NHS Outcomes Framework* for 2013/14 (DoH 2012c) introduced indicators for improvement of maternity care. These were structured in different domains and included: reducing deaths in mothers (or maintaining the low level); reducing deaths in babies (neonatal and stillbirths); provision of help for women recovering from ill health and injury resulting from birth; improving women's

experience of maternity services; treating and caring for people in a safe environment and protecting them from avoidable harm.

2.2.2 MINIMISING COSTS

2.2.2.1 HEALTHCARE COSTS

Healthcare is the single largest government service by expenditure in the UK (£111 billion in 2009). Healthcare accounted for 34% of General Government Final Consumption Expenditure (GGFCE), and 8% of Gross Domestic Product (GDP). Rising healthcare costs therefore have a direct impact on the economy[23], and raise concerns about sustainability of the current provision of healthcare (*Public Service Output, Inputs and Productivity: Healthcare 2011, ONS*). Maternity care cost the NHS around £2.6 billion in 2012-13 (House of Commons Public Accounts Committee 2014).

According to a report by the Office for National Statistics (ONS), more than half (55.8%) of the total £110.6 billion publicly funded healthcare inputs in the UK in 2009 were allocated to labour expenditure (£61.7 billion) (*Public Service Output, Inputs and Productivity: Healthcare 2011, ONS*).

The NHS is the largest employer in the UK and Europe and midwives and nurses represent the majority[24] of employees within the organisation. Nurses, midwives and healthcare visitors accounted for a large share of the labour expenditure - £13 billion

[23] The Government allocates funds to the NHS centrally via national taxation.

[24] There were over half a million registered nurses and midwives on the NMC register for England in 2009 (90% of nurses were women; nearly all midwives were women and 131 were males), plus support staff (286,000) to nurses, midwives and doctors of which 146,000 were healthcare assistants.

were spent in 2009 on NHS pay and pre-registration education[25] (*Front Line Care*, Prime Minister's Commission 2010).

There has been a trend of reduced productivity in the health sector. Productivity fell by 2.7% between 1995 and 2009[26], which represented a decline on average of 0.2% per year. This was mainly due to healthcare inputs growing faster than outputs (ONS 2010).

Several drivers could contribute to rising workload intensity in maternity, which would exacerbate the problem of staffing costs. These drivers include: the public's higher expectations for safety, quality and efficiency from their healthcare providers; the national policies towards woman centeredness (choice, one-to-one care, continuity of care); the provision of care closer to home in the community and policies related to minimum staffing requirements.

These expectations and policy changes are accompanied by tightening of the public spending and less resources. NHS efficiency savings of £20 billion were planned by the end of 2014 (QIPP, DoH 2010) which were to be achieved through improved quality, productivity, innovations, and prevention (QIPP). The QIPP initiative aimed to identify drivers of efficiency and ways of service redesign in order to achieve both improved quality and efficiency.

The prolonged recession, the policy of public funding cuts in conjunction with strategies to improve productivity, quality and safety through reorganisation of the NHS are likely to have a direct and indirect impact on the midwifery and nursing workforce because of the labour intensity of the health sector.

---

[25] There were 20,600 commissioned student places for nurses and midwives in England in 2009.

[26] Positive growth was registered in 2001, 2006 and 2009.

Innovations in healthcare services can contribute to patient satisfaction and better outcomes (Caird et al. 2010). To achieve better outcomes at lower costs, the sustainable answer is *"radical innovation"* in public services, according to Albury (2009). This type of innovation was needed due to pressures created by the combination of increased demands on the public services (e.g. raised expectations); pressing long-term challenges (e.g. ageing population); persistent unresolved issues (e.g. drug and alcohol abuse) and recession (e.g. fewer resources).

Innovation was highlighted by the *High Quality Care for All* (DoH 2008) as a way of achieving efficiency and quality with limited resources. New roles, role substitution and delegation of tasks and responsibilities were seen as some of the innovative approaches in midwifery and nursing. It was recognised that the evolution of midwifery-led models of care was only possible within a workplace culture of mutual respect, which acknowledges professional interdependence. Applying innovative approaches to care requires strong leadership, peer support, creativity and confidence. In midwifery, innovation might be influenced by how risk was managed in complying with specific risk standards (see Clinical Negligence Scheme for Trusts) as these could be limiting to innovation. How risk was managed might depend to an extent on the quantity and the quality of the workforce, the skill mix, deployment, experience, training, leadership and the culture of the workplace.

2.2.2.2 OBSTETRIC LITIGATION COSTS

Maternity services have been and still are associated with much higher litigation costs compared to other services. Maternity related clinical negligence claims are the highest in values and second highest in number in all claims reported to NHS Litigation Authority (2012) and under the CNST. Obstetrics and gynaecology claims accounted for 20% of the number[27] of all the clinical negligence claims with NHSLA and 49% of the total value (1995-2011).

[27] Surgery had the highest number of claims, followed by O&G; O&G had the highest value.

The amount paid by NHSLA between 1995 and 2011 was £5.2 billion (NHSLA 2012). There were 5,087 maternity claims for the period of 2000-2010, with a total value of £3.1 billion, compared with a total of 5.5 million births in England between 2000 and 2009 (NHSLA 2012). The three most frequent categories of claim (2000-2010) were related to management of labour (14%); caesarean section (13.2%) and cerebral palsy (10.6%), (NHSLA 2012). Cerebral palsy and management of labour, together with CTG interpretation accounted for 70% of the total value of all the maternity claims. For example, the cost of lawsuits regarding misinterpretation by a midwife or obstetrician of a fetal heart scan cardiotocogram (CTG) was 11.8m in 2006, rising more than seven times to £85.5m in 2010. There were 130 such cases between 2006 and 2010 and the total amount paid was £196.8m (The Guardian, 14 April 2011). Some 78 babies died and 42 developed cerebral palsy during the same period.

The number and cost of claims are reported to the NHS Litigation Authority. In response to the substantial increase in litigation, the Clinical Negligence Scheme for Trusts (CNST) in England introduced maternity clinical risk management standards (NHS Litigation Authority 2007). CNST is administered by the NHS Litigation Authority (NHSLA), established in 1995 as a Special Health Authority. Membership to CNST is voluntary, but all NHS organisations in England providing maternity services are members (NHSLA 2012), and funding is on "*pay-as-you-go non-profit basis"* (CNST 2012:4). The funds are used to help cover the costs of litigation. Member organisations showing compliance with the maternity standards in CNST receive a discount. The standards aimed to improve safety for women and their babies; provide a framework for risk management activities; improve safety and quality of patient care; assist trusts in identifying risk; encourage incident-reporting and reduction in incident severity; and promote learning from claims. The standards are updated each year, when successfully controlled risks are removed and emerging risks added. There are three levels of compliance with the standards, with respectively 10%, 20% and 30% discounts. There are five separate standards, each

with 10 criteria[28]. The standards apply to the organisation including its staffing levels and skill mix in all settings; clinical care; high risk conditions; communication; and postnatal and newborn care.

The Confidential Enquiry into Maternal Deaths (CEMD) was established in the UK in 1952 as a system of confidential enquiries into the main causes of maternal deaths aiming at providing recommendations for better clinical care and service provision. A system of confidential enquiries into causes of stillbirth and infant death was established in 1992 – the Confidential Enquiries into Stillbirth and Deaths in Infancy (CESDI). It aimed to investigate poor practice and service provision and reduce mortality. Both systems were replaced in 2003 by the Confidential Enquiry into Maternal and Child Health (CEMACH). The Centre for Maternal and Child Enquiries (CMACE) conducted these enquiries during the triennial period 2006-2008, commissioned by the National Patient Safety Agency. Responsibility for overseeing cases of maternal, stillbirth and infant deaths is now the responsibility of MBRACE-UK (since January 2013), hosted and led by the National Perinatal Epidemiology Unit at the University of Oxford and appointed by the Healthcare Quality Improvement Partnership (HQIP). The programme of work is now called the Maternal, Newborn and Infant Clinical Outcome Review Programme (MNI-CORP).

---

[28] **Organisation criteria:** risk management strategy (for organisation and leadership); staffing levels (nurses and midwives, obstetricians and anaesthetists); labour ward staffing; maternity records; incidents, complaints and claims; training needs analysis; skills and drills. **Clinical care criteria:** care of women in labour; use of oxytocin, caesarean section, induction of labour and VBAC; intermittent auscultation, continuous fetal monitoring and fetal blood sampling; severely ill women and high dependency care. **High risk conditions criteria:** eclampsia and severe pre-eclampsia; multiple pregnancy and birth; pre-existing diabetes and obesity; operative vaginal delivery, perineal trauma, shoulder dystocia, PPH; and venous thromboembolism. **Communication criteria:** booking and missed appointments; patient information, antenatal clinical risk assessment and maternal screening tests; mental health; labour clinical risk assessment; onsite handover of care; maternal transfer by ambulance and non-obstetric emergency care. **Postnatal and new-born care criteria:** referral in cases of fetal abnormality; newborn life support; admission to neonatal unit; immediate care of the newborn; newborn feeding; examination of the newborn; bladder care; support for parents; postnatal care and recovery. **Source:** p.32 in CNST Maternity, Version 1 2012/13.

The CMACE (2011) report found that sepsis was the most common cause of direct maternal death and cardiac disease the most common cause of indirect death. The report found that substandard care was a factor in 70% of direct deaths and 55% of indirect deaths. Table 2.2-1 below shows the number of maternal deaths by mode of delivery in 2006-08. 59% of women who died of direct and indirect cause delivered by CS (of them 22% by emergency CS). The report stated that it was difficult to distinguish between cause and effect for nearly all of these women, as most of them had serious prenatal and intrapartum complications or illness and most of the operations were performed to save either the woman's or the baby's life.

Table 2.2-1 Number of maternal deaths by mode of delivery at 24 or more completed weeks of gestation; UK: 2006-08

Mode of delivery	Direct	Indirect	Direct and Indirect		Coincidental	Late Direct	All Deaths	
	N	N	N	%	N	N	N	%
Unassisted vaginal	26	32	58	35	2	2	62	33
Ventouse	2	5	7	4	1	0	8	4
Forceps	1	2	3	2	0	0	3	2
Caesarean section	50	47	97	59	14	5	116	61
Emergency	21	15	36	22	3	1	40	21
Urgent	6	7	13	8	1	1	15	8
Scheduled	3	2	5	3	1	1	7	4
Elective	1	5	6	4	5	2	13	7
Peri- or post-mortem	19	17	36	22	4	0	40	21
Not known	0	1	1	1	0	0	1	1
Total delivered	79	86	165	100	17	7	189	100

Source: Table 1.9, p.40 in CMACE (2011)

2.3 POLICY SUPPORTING NORMAL BIRTH

Normal birth policy refers to a renewed understanding that pregnancy, labour and birth are normal life processes. Increasingly, birth for the majority of women who are healthy is seen as a normal physiological event and not a medical procedure (DoH 2004).

The World Health Organisation first addressed normal birth in a report published in 1996 (WHO 1996). It reviewed the evidence of most commonly used practices (useful, harmful, inappropriately used and those with insufficient evidence) throughout labour and provided recommendations for best practices which support normal birth (irrespective of place of birth or level of care). The WHO definition of normal birth was based on two considerations: the risk status of a pregnancy and continuous assessment during labour and delivery. Given that no risk assessment had 100% predictive power they defined normal birth as: "*spontaneous in onset, low-risk at the start of labour and remaining so throughout labour and delivery. The infant is born spontaneously in the vertex position between 37 and 42 completed weeks of pregnancy. After birth mother and infant are in good condition. However, as the labour and delivery of many high-risk pregnant women have a normal course, a number of the recommendations in this paper also apply to the care of these women"* (WHO 1996:4). The aim of the care provided in normal birth was to achieve a healthy mother and child with the least possible level of intervention that is compatible with safety (WHO 1996:4). Research also showed that most women preferred no interventions as long as their baby and own safety was not compromised (Thomas et al. 2001 and Greene et al. 2003) and that most women wanted spontaneous vaginal delivery (Patel and Murphy 2004).

In England, midwives were seen as experts in the management of normal pregnancy and birth in the policy papers by the *National Service Framework for Maternity Care in England* (NSF, DoH 2004) *and Maternity Matters* (DoH 2007). The NSF (DoH 2004) in England also recognised that *"for the majority of women, pregnancy and childbirth are normal life events requiring minimal intervention"*. The definition of 'normal' labour and birth in England was endorsed by the Maternity Care Working Party (MCWP 2007) in their consensus statement. It used the definition adopted by the NHS Information Centre[29] based on a set of routinely collected statistics and

[29] NHS IC uses the term 'normal delivery' in its maternity statistics publications and considers 'normal delivery' not an outcome but a process of labour measurement.

defined as birth without induction, the use of instruments, CS, episiotomy and without general, spinal or epidural anaesthetic before or during delivery (MCWP 2007). The consensus statement was developed by members of the MCWP (2007) which was an independent, multidisciplinary body[30] dedicated to improvement of maternity care and established to advance the understanding and health implications of the rising trends in caesarean rates. It aimed to promote the positive aspects of 'normal' labour and birth to providers of maternity care and to offer a standard definition of 'normal' labour and birth in order to facilitate auditing, monitoring and comparison of similar women across healthcare providers and models of care (MCWP 2007).

The working party recommended that maternity services should aim to achieve 60% normal birth rate by 2010, supported by strategies which provide women with positive experiences of birth. Other recommendations included providing one-to-one midwifery support for all women in established labour, with appropriate training and staffing levels to achieve that (1 to 1.4 WTE midwives per woman in labour depending on case-mix as recommended by Royal Colleges (RCOG 2007)). In addition, women should have antenatal guidance to acquire alternative skills for coping with pain in labour, and evidence-based accessible information related to the benefits of normal birth and choice of place of birth. MCWP (2007) also recommended that women were given the opportunity to know their midwife prior to labour. Other recommendations and targets included: implementation of the available NICE guidelines related to induction of labour, intrapartum care, caesarean section and fetal monitoring; comparative rates of normal birth across different birth settings; revision of payment by result (PbR) to eliminate the perverse incentives for high intervention rates; commissioning of research to establish the effects of case-mix and

[30] Members of MCWP included representatives from academia, medical professions and organisations such as the Royal College of Midwives (RCM), Royal College of Obstetricians and Gynaecologists (RCOG), Nursing and Midwifery Council (NMC), Centre for Research in Midwifery and Childbirth (CeMaC), the National Childbirth Trusts (NCT), BirthChoiceUK, BirthCentre Network UK, Birth Crisis Network, Independent Midwives Association, Association of Radical Midwives.

organisational factors on normal birth rates; and annual publication of normal birth rates statistics using the NHS IC definition in all four countries in UK. The measurement and audit of normal birth rates were supported by the Department of Health (2009) and the Royal Colleges (RCOG 2008).

The normal birth rate in England in 2010-11 was 42%, with unadjusted rates in hospitals ranging from 34% to 49% (BirthChoiceUK). Various studies have identified factors which supported normal birth, such as: one-to-one midwifery care (Sandall et al. 2013); better training of junior doctors and known midwife (Spencer 2006); support from consultant obstetrician and involvement of consultant midwife (NHS Institute 2006); immersion in water (Cluett et al. 2004); for low risk women planning a home birth (NICE 2004); and significantly higher odds of normal birth for low risk women in all three non-obstetric settings (AMU, FMU and home births, Birthplace 2011).

The culture of the NHS organisation could impact on normal birth rates, for example regarding the degree of positive outlook on birth as a normal physiological process, as well as support for junior staff, promotion of evidence-based practice, leadership, integration between different services and good communication (MCWP 2007).

Dodwell and Newburn (NCT 2010) recommended normal birth as a measure of the quality of the overall maternity care and specifically of midwifery care. They argued that normal birth indicator was a good measure of quality based on the quality definition of safe, effective and patient-centred care, and was beneficial in providing positive birthing experiences for women.

2.4 CONTINUITY OF CARE

A report published in 1992 by the House of Commons Health Committee chaired by Sir Nicholas Winterton recommended that midwives should take full responsibility for women in their care by independently carrying their own caseload and through

midwifery managed maternity units within and outside hospitals (House of Commons 1992).

Similar recommendations regarding midwifery roles and provision of greater choice for women, including place of birth and professional providing care, were highlighted in *Changing Childbirth* report by the Expert Maternity Group (DoH 1993) the following year. The report recommended that every woman should have continuity of care from a known midwife, and should be familiar with the lead professional in control of her care.

These reports recognised the increased professional independence of midwives in Britain (WHO 1997); however implementation of recommendations was problematic. Ten years later the House of Commons Select Committee (2003) and the independent think-tank Reform (2005) which reviewed whether the *Changing Childbirth* recommendations have been followed, found that relatively few had been implemented.

The RCOG and RCM joint report *Towards Safer Childbirth* in 1999 supported the principles of midwifery care; recognised the autonomous position of midwives providing care in normal labour and birth and redefined the role of the consultant obstetrician[31]; In 2004 the *National Service Framework (NSF) for Children, Young People and Maternity Services*, and *Maternity Matters* published in 2007 supported midwife-led care for women, and providers were encouraged to develop midwife and home birth services in response to local needs. However, strong evidence regarding safety of place of birth did not exist at the time to support these initiatives. A joint Royal Colleges[32] report in 2007 *Safer Childbirth* acknowledged the autonomous role of midwives providing care in normal labour, as well as their partner role within a multidisciplinary team providing care in complex pregnancies and labour. Primarily

[31] By accepting regular responsibilities for labour wards as well.

[32] RCOG, RCM, RCA, RCPCH.

though, the report was published due to persistent concerns related to inadequate staffing levels in all the professional groups involved (obstetricians, junior obstetric staff, midwives, maternity care assistants, anaesthetists and paediatricians). It recognised that a considerable expansion in numbers of midwifery and medical staff was needed and recommended minimum standards for the organisation and delivery of care in labour related to organisation, multidisciplinary working, staffing levels, leadership, responsibilities, emergencies and transfers, training and education, facilities, and childbirth outcomes.

2.4.1 WOMAN-CENTRED CARE

A team from the National Institute for Health and Care Excellence (NICE – formally the National Institute for Health and Clinical Excellence) conducted a systematic review of 137 studies to establish which factors most influenced women experiences of childbirth (NICE intrapartum guideline 2007).

Four factors were identified as key in women's experiences of labour: personal expectations; the amount of support from caregivers; the quality of the caregiver-patient relationship; and the involvement in decision making. The attitudes and behaviours of the caregivers appeared to be more powerful in predicting women satisfaction than the individual experience of pain, provision of pain relief or intrapartum interventions.

The NICE intrapartum guideline recommended that all women and their partners should be treated by their healthcare professionals with respect, kindness and dignity; that establishing good communication and providing them with evidence-based information was essential. "*The views, beliefs and values of the woman, her partner and her family in relation to her care and that of her baby should be sought and respected at all times"*(NICE 2007:5).

Continuity of care throughout pregnancy and childbirth was seen as an important component of women-centred maternity care and *"there is evidence that continuity*

of care in complex organisations may be associated with increased patient safety" (DoH 2005:203). *Transforming Community Services* initiative (DoH 2009) aims to enable midwives to provide continuity of care from their first contact with women in the antenatal period through the last postnatal contact.

The integration of midwifery services in maternity networks was seen as a way to promote continuity of care (CfWI 2012; DoH 2010a; RCOG 2011). Maternity networks were also seen as a way to share good practice and reduce variations in outcomes. These were intended to bring together commissioners, providers and other stakeholders (including users of maternity services) with the aim of achieving the best outcomes for women and babies. The networks were not compulsory and the ones that currently exist are not as well developed, as other NHS networks (i.e. neonatal networks). Currently 25% of trusts in England are not part of a network (House of Commons Committee 2014).

2.4.1.1 WOMEN'S CHOICE OF PLACE OF BIRTH

The policy of choice for all women, including the choice of place to give birth, is another key policy factor, as established by the *National Service Framework (NSF) for Children, Young People and Maternity Services* (DoH 2004) (See Box 1 in the Appendix).

The *'national choice guarantee'* policy by Department of Health (*Maternity Matters*, DoH 2007) stated that by the end of 2009, all women depending on their circumstances should be offered choice of place of birth; type of antenatal and postnatal care; and choice of how to access maternity care. The provision of choice antenatally related to women choosing the best antenatal pathway of care, such as location, and number and time of appointments, including provision outside normal working hours. Choices during labour related to place of birth (obstetric or midwife-led unit) or home births; interventions during labour (and the nature of the professional support) and choice of pain relief. It was expected that these choices would be sustained through promotion of normal birth and changes to the

organisation of midwife-led care. Postnatally the choices related to number of contacts women had with a midwife, maternity support workers, clinics or by phone to deal with issues of breastfeeding, skin-to-skin contact, crying, sleeping, and general wellbeing of mother and baby. To support the best possible care and a full range of birthing choices, the Government in 2008 announced extra funding for maternity, totalling £330 million over the following three years.

Towards Better Births (HCC 2008), a review of maternity services in England conducted by the Commission for Healthcare Audit and Inspection, revealed some worrying deficiencies in maternity services, such as: inadequate staffing levels and continuity of care; not enough consultant cover on the labour ward; non attendance of training by doctors and midwives; and poor postnatal care. The review also found that 80% (of 26,325) women were given the choice of where to have their baby at the start of pregnancy and 58% were offered home birth. However, only half of the women reported being offered sufficient information on which to base this choice. Choice was limited by the lack of midwifery-led units (65% of the then 152 trusts in England reported having only obstetric units). The remaining trusts had a combination of obstetric and midwife-led units (either alongside the main unit or freestanding on separate premises) and only few had all three.

The *Birthplace* in England study (2011) compared the configuration of maternity services in 2010 with 2007 and found that choice of place of birth had increased, mainly via an increase in number or capacity of alongside-midwifery units in some regions, but still around 50% of trusts had only obstetric units. While options for place of birth in 2010 had improved, the study also revealed that about half of all women did not have a full range of choices; only 10% actually gave birth outside of an obstetric-led unit; and there were regional variations and inequalities in type of services provided. The study recognised shortages of midwives and warned that an increase in home births and midwifery units was a labour intensive option which would require more midwives and forward planning. Comparing the costs of midwifery and obstetric units, staff ratios were higher in the former, but intrapartum care costs were higher in the latter. It was found that obstetric units provided less

one-to-one care than other settings. Expanding midwifery units would provide more one-to-one care, requiring more midwives.

The *Birthplace* study in England[33] (2011) aimed to fill the gap in evidence related to the configuration of maternity services; availability of different models of care across providers; comparison of maternal and infant outcomes between settings (safety of planned place of birth); and cost-effectiveness of different settings all in relation to the quality and safety of maternity care. The *Birthplace* study results confirmed that overall giving birth in England was very safe and that offering healthy women with low risk pregnancies a choice of place of birth was a safe policy. Some of the other findings were that giving birth in a midwifery-led unit was a safe and cost-effective alternative to obstetric units for low-risk women with the added benefit of fewer interventions (around half the rate of caesarean section for low risk women). Outcomes for babies were the same in midwifery and obstetric units. However there was an increased risk for the baby when women had their first baby at home; while home births, for women having a subsequent baby, did not increase the risks for the baby.

The current Coalition Government announced after its election a long-term plan in the white paper *Equity and Excellence Liberating the NHS* (DoH 2010). They expressed their commitment to women's choice (DoH 2012a and DoH 2012b) and focus on achieving better outcomes (DoH 2012c), but have not published yet a specific maternity policy. That commitment aimed to improve safe, informed choices in pregnancy and childbirth by developing new provider networks, while recognising that not all choices would be appropriate or safe for all women.

[33] The Birthplace Research Programme was commissioned in 2007 and executed by the National Perinatal Epidemiology Unit (Oxford). It was funded by the NIHR Service Delivery and Organisation Programme and DoH Policy Research Programme.

2.4.2 MIDWIFE-LED CONTINUITY OF CARE

Midwifery models of care are based on a particular view of pregnancy and birth which differs from the medical model, in that normality and the natural ability of women to experience birth with minimum interventions is acknowledged and encouraged (Hatem et al. 2008). Midwifery models also differ in their objectives of care, in the nature of the relationship between the care provider and the woman; in philosophy on interventions during labour and care settings (Rooks 1999).

Midwifery care recognises that the pregnant woman is the decision-maker, the main actor and an active partner; her and her family views, beliefs and values should be respected at all times; the midwife's role apart from being competent practitioner of normal labour and birth is to provide information, advice, to encourage, empower and comfort women and to support them in their informed decisions through the antenatal, intrapartum and postnatal stages and that interfering with 'normal' processes is to be avoided.[34] During labour, 'normal' has a narrower definition in the medical model of care while the midwifery model of care accepts a wider definition of 'normal' as long as the mother and foetus are coping well in labour. The midwives involvement in women's lives was recognised as time-intensive, labour-intensive and relationship-intensive (Rooks 1999).

The term midwife-led care (Hatem et al. 2008) seemed to have been replaced by midwife-led continuity of care in the updated Cochrane review (Sandall et al. 2013) while retaining the same definition "*where the midwife is the lead professional in the planning, organisation and delivery of care given to a woman from initial booking to the postnatal period"* (RCOG 2001 cited in Sandall et al. 2013:2). The midwife may interact with other professions (make referrals) at each stage (antenatal, intrapartum, postnatal), but she remains the responsible, lead professional in assessing, planning,

[34] A key priority for implementation in NICE 2007 guidelines in regard to 'normal labour' stated: "Clinical intervention should not be offered or advised where labour is progressing normally and the woman and baby are well".

organising and delivering care to women. Midwife-led care is provided in hospitals or community setting, usually to healthy low risk pregnancies, but could be provided to all women in a defined location, so that midwives remain the lead professional for women with uncomplicated pregnancies while continue also to provide midwifery care to women with complex needs in partnership with other professionals.

Midwife-led continuity of care was reviewed by Sandall et al. (2013) in their update of the Cochrane review (Hatem et al. 2008). The authors reviewed 13 randomized control trials in which midwife-led continuity of care was compared to medical and shared models of care and identified several benefits for mother and baby and no adverse effects of midwife-led continuity of care model. The benefits included: reduction in the use of epidurals, fewer episiotomies or instrumental births; increased chance of spontaneous vaginal birth; and no difference in the number of caesarean births; women were less likely to experience pre-term birth or to lose their baby before 24 weeks gestation; they also had increased chances to be cared in labour by a midwife they had got to know.

The term continuity of care in maternity services relates to both continuity of carer and consistency of care (NICE 2007). In most policy reports and research though, continuity of care was identified with continuity of carer concept (NICE 2007). Continuity was defined by Freeman (2007 cited in Hatem et al. 2009) as having three dimensions: management, information and relationship. Relationship continuity (patient-health professional relationship over time) was seen as having a great impact on patient experiences and outcomes (Saultz 2005 cited in Hatem et al. 2009).

Continuity of carer relates to the provision of care from a midwife or a small group of midwives through the antenatal, intrapartum and postnatal periods. It was seen as a key concept in provision of good maternity care in the *Winterton Report* (House of Commons 1992) and *Changing Childbirth Report* (DoH 1993) (see page 10). On a related matter, women should have access to a named midwife, according to various policy papers. *Changing Childbirth* (DoH 1993) recommended a named midwife for every woman who would be responsible for providing continuity of care. The *Front*

Line Care report (2010) stated that services should be organised "*so that every woman has a named midwife responsible for ensuring coordination of her care and providing support and guidance".*

In order to provide continuity of carer within the NHS structure and in a sustainable way, in the 1990s several models of midwifery practice were piloted which allowed women to know the midwife who would care for them during delivery. The two main models were team midwifery and caseload midwifery. In both models a team of midwives provided care to a group of women; however caseload midwifery teams were smaller (usually 2 midwives) and aimed to establish more personal relationship with the woman.

2.4.2.1 TEAM MIDWIFERY

Team midwifery was defined by NICE (2007) as "*a group of midwives providing care and taking shared responsibility for a group of women from the antenatal, through intrapartum to the postnatal period".* Teams were based in the community, in hospitals, or integrated across. The team sizes varied between four to ten or more midwives, with hospital-based teams being larger. Organising care in such a way was labour intensive and required midwives to be constantly available for delivery[35] which was a huge commitment for most of them who had families of their own. Though many women recognised the value of continuity of carer there were concerns about midwives burnout (particularly hospital-based), working under continuity of care system (Sandall 1997; Sandall 1998).

A meta-analysis included in the NICE intrapartum guideline (2007) suggested that women who received antenatal, intrapartum and postnatal care from a team of midwives had fewer interventions during labour and birth (lower rate of instrumental

[35] Many midwives do not like to say that they *"deliver babies"*; they prefer to *"attend" the labouring woman* and *"catch" the baby*, in acknowledgment that a woman herself, through labour, delivers her own child. Physicians are more likely to say that they "deliver" babies as they see themselves to make the decisions and to be in charge (Rooks 1999).

vaginal births). However there was evidence of increased perinatal mortality associated with team midwifery care but little understanding of what contributed to it. There was little evidence of the cost-effectiveness of the scheme. The service though was perceived by providers as costly, resource-intensive and less effective compared to the conventional models of care (though some potential for savings was seen in the fact that reduced rates of instrumental vaginal births might be able to offset higher staffing costs); thus long-term funding support for the scheme was not provided (NICE 2007). For these reasons team midwifery was not widely established and most of the teams ceased to exist. Team midwifery was also not recommended for care during labour (NICE clinical guideline 55, 2007)

2.4.2.2 CASELOAD MIDWIFERY

Caseload midwifery is defined as *"system of care whereby one midwife (sometimes referred to as the 'named midwife') is responsible, and provides the majority of the care, for a group of women backed up by a small group of associate midwives (usually two or three)"*(NICE 2007). When the named midwife was backed up by another midwife, it was sometimes called a one-to-one system of care (NICE 2007). Caseload midwifery teams are mostly community-based.

The NICE intrapartum guideline (NICE 2007) outlined the need for future studies to investigate the cost-effectiveness of the model and the impact of caseload midwifery on long-term outcomes, on women's experiences and on the workforce.

Other innovative schemes exist and have been investigated – these included a package of care which aimed at providing continuity of care in diverse settings – midwife-led units, traditional delivery suits, birthing rooms within midwifery suite, separate birthing units.

2.4.3 ONE-TO-ONE CARE

NICE defined one-to-one care as:*"continuous presence and support either by husband/partners, midwives or other birth supporters during labour and childbirth"*

(NICE 2007). Page (2003) defined one-to-one midwifery as a model of care developed in the UK, aiming to provide a continuous personal relationship between a woman and her midwife. She also suggested that one-to-one type practices were well suited to serve communities with diverse ethnic and socially and economically disadvantaged populations. Many women in the UK give birth with support from their partners who are committed to meet some of these needs.

NICE (2007) suggested that the support women needed was four dimensional: emotional, physical, information and advocacy. *Maternity Matters* (2008) recommended that all women in established labour should receive one-to-one care from a midwife ("*as far as practicable*"). NICE recommended that women in established labour should receive one-to-one supportive care; they should not be left alone, except for short periods or by request; and that women should be given a choice of birth partner (NICE 2007). NICE (2007) also examined other evidence for one-to-one care benefits and found that the risks of instrumental vaginal birth and caesarean section were significantly reduced; women had more positive experiences of childbirth and higher satisfaction; and that these benefits remained if the support was provided by a non-professional staff member. Evidence to compare one-to-one support from midwives with other professionals was not identified. The evidence from the benefits of lay person support was from countries with different maternity care settings to UK and it was considered inappropriate to extrapolate these findings to UK where midwives provide the majority of care in labour. The NICE Caesarean section clinical guideline (2004) also recommended that women should be informed about the benefit of continuous support during labour as the likelihood of caesarean section was reduced when women received continuous support during labour from clinically trained or untrained women.

2.5 ORGANISATION OF MATERNITY SERVICES

Organisation refers to the distribution of maternity services across and within hospitals.

Since 1973, there has been increased centralisation of maternity services in England. There were 527 maternity units in 1973 in England and 341 in 1996. The justification for centralisation included cost cutting and improving patient safety (REFORM 2005). One of the consequences however is reduced choice for women and increased workload for maternity staff.

The *Birthplace* in England (Redshaw et al. 2011) research programme provided findings from the mapping of maternity care in 2007 and 2010. Prior to this study there was little reliable information on the distribution of midwifery-led units (freestanding midwifery units-FMUs and alongside midwifery units-AMUs), their geographical location and relationship to obstetric and home birth services. The data from 2007 were collected by a mandatory survey which was carried out as part of the Healthcare Commission review of maternity services *Towards Better Births* (HCC 2008). The second survey was carried out in 2010 by the *Birthplace* research team but only 63% (93/148) of the trusts responded (these were representative in terms of configuration); data related to trusts configuration and type of units was available for 100% of the trusts. In 2007, there were 152 trusts with maternity services, of which 66% had only obstetric units (1 or more); this proportion decreased to 49% by 2010. The overall number of maternity units had increased over that period by 11% (mainly due to AMUs which doubled from 26 in 2007 to 53 in 2010). There were wide variations in the number of midwifery-led units across England in 2007 (FMU were most common in the South West, while AMU in London and South Central SHA); and inconsistencies regarding eligibility criteria for admission alongside or freestanding midwifery units. There were also wide variations in midwifery staffing levels (per 1000 births) across units of the same type and across different type of

units[36]; similar variations were observed regarding obstetric medical staffing levels in obstetric units (median of 6.8 per 1000 women).

The most recent data regarding the number of maternity units in England was provided by BirthChoiceUK and the National Audit Office (2013), comparing 2007 to 2013. There were 268 maternity units in 2007 (181 obstetric units, 31 alongside midwifery units and 56 freestanding midwifery units). By 2013 the total number of maternity units were 316 due to a significant increase in midwifery-led units (alongside and freestanding) from 87 in 2007 to 152 in 2013; the obstetric units have decreased over that same period from 181 in 2007 to 164 in 2013. Despite the increased number of midwifery-led units and an increase in proportion of births in such units only 11% of women gave birth in a midwife-led unit in 2012 (87% gave birth in an obstetric unit and 2.4% were home births, NAO 2013), possibly because midwifery units are usually much smaller than obstetric units and therefore most of the births continue to occur in obstetric units (Birthplace 2010).

The definitions[37] of different types of maternity units (obstetric/alongside/freestanding) were provided by the *Birthplace* in England (2011) study:

- Obstetric unit (OU) – this is an NHS clinical location where care is provided by a team; obstetricians take primary responsibility for high-risk women, while midwives take primary responsibility for low-risk women, but still care for all women admitted. Obstetric, neonatal and anaesthetic services are available on site.
- Alongside midwifery unit (AMU) – this is an NHS clinical location, co-located in the same building or site as an obstetric unit. Midwives are the lead

[36] Median of 35 midwives per 1000 women giving birth in a FMU and 31 per 1000 in AMUs and OUs.

[37] Birthplace in England Research Programme, 2010 Update: Configuration of Maternity Care Unit Questionnaire, 2010.

professionals and take the primary responsibility for care of low risk women. Obstetric, neonatal and anaesthetic services are available on site if needed. Transfer of women is provided by wheelchair or trolley/bed.

- Freestanding midwifery unit (FMU) – this is an NHS clinical location. These units are sometimes called birth centres and are geographically separate from hospital obstetric or consultant-led units. Midwives are the lead professionals and take the primary responsibility for care of low risk women. Obstetric, neonatal and anaesthetic services are not available on site and transfer of women is provided by car or ambulance.
- Home births services – community midwives provide labour care at women's home.

2.6 DEMAND-SIDE PRESSURES ON MATERNITY WORKFORCE

2.6.1 BIRTH TRENDS AND MOTHERS' SOCIO-DEMOGRAPHIC CHARACTERISTICS

There has been a steady upward birth trend in the first twelve years of the new millennium, with birth rates rising to a level not seen since 1972. The ONS statistics show a general upward birth trend since 2001 (Table 2.6-1) with approximately 730,000 live births in 2012, a rise of 23% compared to 2001.

Table 2.6-1: Live births in England, 1972-2012

Year	Live Births (thousands)
1972	725.4
1976	584.3
1981	634.5
1986	661.0
1991	699.2
1996	649.5
2001	594.6
2002	596.1
2003	621.5
2004	639.7
2005	645.8
2006	669.6
2007	690.0
2008	708.7
2009	706.3
2010	723.2
2011	723.9
2012	729.7

Source: ONS, Birth Summary Tables, England and Wales 2012

The increase in births has been attributed to the increasing fertility rates among women born in the UK, and an increase in the population of childbearing women not born in the UK. A quarter (25.9%) of births in England and Wales in 2012 were to mothers born outside the UK, compared to 11.6% in 1990. The last issue was particularly prominent for London where 57.4% of women giving birth in 2012 were born outside the UK (ONS 2013d). Poland, Pakistan and India were the most common countries of birth for non-UK born mothers in 2012, and Romania moved to

the top 10 most common countries of birth for non-UK born mothers (ONS 2013d). The majority of women of childbearing age though were born in the UK (80% in 2012) and they contributed most to the total fertility rate.

Babies born in 2012 in England and Wales were most likely to be born to women aged 25-34 years (57 % of women were in this group). Nationally around 49% of all live births in 2012 were to mothers aged 30 years and over. The standardised average age of a woman giving birth for the first time in 2012 in England and Wales was 28.1 years, while the average age for all births was 29.8 years (ONS 2013c)[38]. The average age of a mother has been increasing since 1973. Possible reasons included: increased participation in education and the labour market, higher opportunity costs of childbearing and career importance and other economic factors (ONS 2013c). Overall, 84% of births in 2012 were to parents who were married, in a civil partnership or cohabiting.

Teenage live births (under the age of 20 years) have gone down by 26%, while births to women over the age of 40 years have increased by 99% for the period 2000-2012 in England and Wales (ONS 2013e: Table 2a).

Regarding the breakdown of births by parity, ONS estimated that in 2008 around 43% of births were first births; 33% second births; 14% third; and another 9% were fourth or subsequent births (ONS 2009). This distribution has been relatively unchanged since 2001. The largest numbers of first births were to women under 30 years of age. The majority of women were healthy, had a straightforward pregnancy, and most had spontaneous labour onset (NICE 2007). Almost 90% of women gave birth to a single baby after 37 weeks of pregnancy, with the baby presenting head first (NICE 2007 Intrapartum Care guideline). Caesarean births rates in England have

[38] Statistical bulletin: Live Births in England and Wales by Characteristics of Mother 1, 2012.

increased in 2013 to 25.4% from 24.5% in 2009 (BirthChoiceUK 2013 and HSCIC 2013).

2.6.2 RISE IN COMPLEX PREGNANCIES

2.6.2.1 IMPACT ON MATERNITY SERVICES

A report published in 2011 *Improving maternity care in London* by the NHS Commissioning Support for London[39] revealed the complexity of demographic and public health needs faced by maternity services in London. Services were increasingly stretched due to a steady increase in the number of births per year (18% between 2002 and 2008 in London), which led to temporary closure of some units. Provision of continuity of care was also problematic. Women in London gave birth in hospitals located within 0-5 miles of home, with some women giving birth in hospitals outside of their catchment area for their antenatal and postnatal care. The London report (2011) identified several challenges which were expected to stretch maternity services and to increase demand for midwives and specialist staff in London. These challenges applied nationally, with local variations, and so were expected to have an impact on the maternity workforce across NHS trusts in England. A review of those challenges follow (these are based on *Improving Maternity Care in London* 2011 report).

2.6.2.2 TEENAGE MOTHERS AND MOTHERS AGED >40 YEARS

Teenage mothers (age 11-19) and mothers over the age of 40 present challenges: the teenage group in terms of higher risks of stillbirths, and higher rates of perinatal and neonatal deaths (CEMACH 2009); older mothers (>35) with their increased risk of hypertension, gestational diabetes, low lying placenta and the need for more scans and antenatal tests (Bell et al. 2001) as well as continuous antenatal assessments. Older mothers present additional issues related to rise in fertility treatments, multiple

[39] Endorsed by RCM and RCOG.

pregnancies and perinatal loss, and generally increased risk of complications and interventions. This was expected to create higher demand for specialist staff and use of more resources.

2.6.2.3 FERTILITY TREATMENT

In 1992, 14,057 women (average age of 33) in the UK received fertility treatment, while in 2007 this increased to 36,648 (average age of 36) (HFEA 2009). Twin pregnancies are 20 times more likely as a result of fertility treatment and around 25% of such pregnancies result in twin births (HFEA 2009). Multiple pregnancies carry an increased risk for mothers associated with diabetes, hypertension, pre-term labour, post partum haemorrhage and complications during birth. The risks for babies relate to prematurity, low birth weight and admission to neonatal units (NICE 2009). There were 15 per 1000 multiple deliveries in England and Wales in 2007 compared to 9.8 per 1000 in 1980 (NICE 2009). Twin pregnancies demand additional resources and competencies.

2.6.2.4 OBESITY

First trimester maternal obesity in England doubled from 7.6% to 15.6% between 1989 and 2007 (Heslehurst et al. 2010). Obesity is related to miscarriages, diabetes, pre-eclampsia, PPH, wound infections, increased risk of CS and complicated labour. These complex needs create pressure on the provision of antenatal care as they require thorough and continuous risk assessments and multidisciplinary care and significantly increase the cost of antenatal care (CMACE and RCOG, 2010). The centre for Maternal and Child Enquiries and RCOG published joint guidelines related to obesity in 2010. Pre-pregnancy guidance on healthy lifestyle was seen as essential.

Maternal death

Maternal mortality though rare in UK (1 in 20,000: WHO 2007), has not improved in the last 10 years (Hogan et al. 2010). Women from south Asian and Black African communities, asylum seeking women and women living in poverty were more likely to die in childbirth compared to White women in England and Wales (Hogan et al. 2010).

Infant mortality

The rate of infant mortality (IMR, deaths under 1 year old) in England and Wales in 2008 was 4.6 deaths per 1000 live births (ONS 2009). Ethnic variations in infant mortality exist: IMR rate was 6.6 per 1000 for women born in New Commonwealth countries; 9.5 per 1000 for women born in Caribbean countries; 7 per 1000 for women born in East Africa and 8 per 1000 for women born in rest of Africa. The IMR rate for women born in England and Wales was 4.4 per 1000. 68% of the deaths were neonatal deaths (0-28 days old), of them just over 50% were perinatal (0-6 days old); and 31 % were postnatal deaths (28 days to 1 year old). Problems during pregnancy explained half of the postnatal deaths, while 72.5% of infant deaths and 85.8% of neonatal deaths were related to immaturity, congenital abnormalities and antepartum infections.

CEMACH (2007) found a relationship between smoking, alcohol and substance misuse and likelihood of stillbirth and neonatal death as well as a correlation between infant mortality and ethnicity and socio-economic status.

Low infant birth weight (LBW)

Low infant birth weight was defined as weight at birth of less than 2,500g (WHO 2004). '*LBW can be a consequence of preterm birth (defined as birth before 37 completed weeks of gestation), or due to small size for gestational age (SGA, defined*

as weight for gestation <10th percentile), or both' (WHO 2011). The cut-off measurement is for the purposes of international comparison and was established from epidemiological observations that infants weighing less than 2,500g were approximately 20 times more likely to die than heavier babies (WHO 2004). Low birth weight is regarded as major contributor to perinatal mortality and morbidity (diabetes, heart disease, cognitive impairment and reduced immune function). In England and Wales 7.0% (50,918) of all live births were low birthweight (under 2.5kg) in 2012 (ONS 2013c).

Pre-term babies

Pre-term birth is defined as birth before 37 completed weeks of gestation (WHO 2004). Pre-term birth causes low birth weight and is a risk factor for neonatal mortality and morbidity (cerebral palsy, respiratory illness, learning disability). Nutrition, smoking, alcohol/substance misuse, teenage mothers, socially deprived women, women from certain ethnic groups, assisted conception and multiple pregnancies are all linked to pre-term births (Goldenberg 2008).

2.7 SUPPLY-SIDE FACTORS AND STAFFING LEVELS

2.7.1 MIDWIVES

Midwives are involved in the provision of antenatal, labour and postnatal care, underlining the importance of their labour supply. Given the wide range of settings in which they provide care, the multitude of factors affecting shortages of midwives, requirements for continuity of care provision and one-to-one care and higher standards for quality and safety, various reports have found that planning for staffing levels and skill mix had to take into account local models of care, case mix, the needs of women and their families and service design (NMC 2006; RCOG 2007).

Midwives in the UK have long been recognised as autonomous experts in caring for women who have a low risk pregnancy (RCOG 1999). Even after a referral to the obstetrician in complex cases, they are responsible for providing holistic care to

women and promoting normal physiological processes (RCM 2006). They provide antenatal, intrapartum and postnatal care to women in diverse birth settings (at home, in midwifery units and obstetric units within acute hospitals, birth centres and community).

Midwives with appropriate training can complete the newborn physical examination within 72 hours of birth, and can be the lead professional providing care for the baby during elective caesarean section. Some midwives are undertaking new competencies, such as ventouse and vaginal breech births and establishing and leading perineal care clinics. Regardless of place of birth and irrespective of risk category, women and babies in the UK are predominantly cared for by midwives. Midwives practice within multidisciplinary teams, and have complimentary and inter-related responsibilities. They have responsibilities for teaching and mentoring student midwives, and expected to contribute to junior doctors and medical students training. Their roles, responsibilities and accountability are clearly established in statute (NMC 2004) including EU legislation.

The midwifery workforce may comprise student midwives, junior and senior midwives, consultant midwife[40], head of midwifery services[41], labour ward manager[42], labour ward shift coordinators[43]; supervisors of midwives[44]. These

---

[40]'Consultant midwife' role replaced 'clinical midwife lead' role. They have clinical leadership role without managerial responsibilities. They are the experts in clinical midwifery decision-making (referral and transfer of care), promote normality in labour, implement innovative models of care, provide assistance to midwives and medical staff in enhancing their skills for normal labour, promore midwifery research and support provision of safe and effective care (RCOG 2007).

[41] These are experts in women and children's health and provide operational and strategic general management and professional leadership. Responsible for the budget, human resources, strategic planning, clinical governance and quality of midwifery care (RCOG 2007).

[42] They are responsible for resource management (positive ward environment, ensure systems are in place for learning and staff development and for mentoring new and junior midwives), in charge of the evidence-based guidelines, risk management framework to ensure service quality (RCOG 2007).

different roles are assessed by a Job Evaluation Scheme, based on national job profiles and placed in Agenda for Change[45] pay bands 5 to 8. The Agenda for Change was introduced in 2004 and was regarded as the most radical reform of the NHS pay system. These bands were not meant to reflect the job title, rather the level of knowledge, responsibilities, skills and efforts needed for the job (NHS Employers 2009). Newly registered midwives enter at Band 5 and majority of registered midwives are in Band 6. In 2008, 63.7% of midwives were in Band 6; 25.5% were in Band 7; 8.5% were in Band 5 and 0.9% were in Band 8 (*Midwifery 2020*, Workstream 2010).

The RCM (2010a) advocated a national average ratio of 28 births per 1 full-time equivalent (FTE) midwife for midwives working in hospital/labour wards, based on the *Birthrate Plus* tool (Ball et al. 1996, see further for description of the tool). For home births this ratio was 35 births per 1 FTE midwife. The minimum ratio of 1:28 was based on the capacity to achieve one-to-one care in labour while providing a safe level of service (Ball et al. 2003; Greene et al. 1998). An average ratio of 1:25 was suggested (RCOG 2007) for services caring for women with complex needs, and this ratio should be determined by local case-mix (social and clinical determinants) and external workload assessment. These are all total care ratios and indicate the maximum number of women per midwife per annum, based on midwife's antenatal,

---

[43] A rota of experienced senior midwives to ensure 24 hours managerial cover, in charge of staffing numbers required for one-to-one care and communication between professionals (RCOG 2007).

[44] Experienced midwife with additional training in charge of developing maternity strategy in line with safe provision of evidence-based midwifery care. They are advocates for women, midwives and wider multi-professional team. They have a statutory role, undertaken on behalf of local supervising authorities and integral part of the clinical governance framework (participants in all lines of communication and forums: maternity services liaison committees, risk management, labour ward forums, perinatal audit meetings and trust executives). Available 24 hours. (RCOG 2007).

[45] AfC has 9 pay bands.

intrapartum and postnatal care within the service and it does not consider all midwifery roles.

The midwife-to-live-birth ratio in England in 2001 was 1:33 (18,048 FTE midwives), ten years later in 2010 it was 1:35 (20,790 midwives) (CfWI 2012). Despite an observed increase in midwives FTE, the change in ratio over time reflected the increased birth rate in England. In any case, it indicated deviation from the recommended midwife-to-woman ratios and suggested the need for expansion of the midwifery workforce to align with the rising birth rate and possibly revaluating the best use of midwife time. The use of more support workers were seen as facilitating this, mainly through performing routine tasks, such as taking blood samples, making observations – temperature, pulse, blood pressure, carrying out administrative duties, preparing equipment and cleaning up after sessions (CfWI 2012; NHS WRT 2009).

The midwifery profession has expressed concern about the increase in complexity[46] of cases, especially in London, and about the local staffing needed to meet not only an increase in local birth rates, but a rising complexity of cases, provision of different models of care and local services configuration. In addition inadequate staffing levels were seen by heads of midwifery services (RCM 2005) as affecting negatively midwives attendance to additional training and professional development as well as clinical support to inexperienced staff and students and direct care, due to more clerical work and the contingency plans for unexpected surges in births.

[46] In emergencies, the higher dependency of women in labour will create a fluctuating demand for extra staff and resources, which can stretch the system if was not anticipated or planned for in advance.

2.7.1.1 BIRTHRATE PLUS (BR+)

Birthrate (Ball 1989) and *Birthrate Plus*[47] (Ball et al. 1996) are evidence-based staffing planning tools used to forecast midwifery numbers required on a labour ward. *Birthrate* was approved by the Royal Colleges (RCM; RCOG) which recommended the tool to the House of Commons Select Committee on Maternity Care in 1991 and *Birthrate Plus* was endorsed by the Department of Health in England as the workforce planning tool in maternity services (RCOG 2007).

BR+ is based on the underlying principle that labouring women should receive one-to-one individual care from midwives in established labour. It used clinical indicators to categorise women into five case-mix classifications and measured midwifery time needed to provide one-to-one care. The tool measures the volume of workload adjusted for each of the five case-mix classifications; it is sensitive to the intensity of workload demands and models of care and has the capacity to inform staff deployment and recommend required midwifery staffing. It allows for local adjustments to take into account not only the NICE quality standards but local needs; initiatives and knowledge. At local level, it is used for calculating intrapartum staffing ratios based on all expected annual number of births by simultaneously accounting for case mix and midwifery skill mix. However it takes into account midwives contribution but not contributions from other staff (such as midwifery managers, consultant midwives).

The *Birthrate* tool had three main components: score system, midwife hours and staffing formula. The original planning tool was refined and consequently replaced by *Birthrate Plus*. It is widely in use as it adjusts the required staffing numbers to the individual labour ward's demand, case mix and for specific care settings. The focus of BR+ was women's needs, not midwifery activity. BR+ uses a retrospective score

[47] The term *Birthrate Plus* is a registered trademark of Birthrate Plus Consultancy Ltd.

system allocated according to key clinical indicators of need for midwifery care during labour and delivery for both mother and baby. The scores, based on need, are then summed up and allocated in five categories – categories I and II being the most normal and category V representing the highest level of need. The assessment of midwife hours to match the needs in each category was based on the length of time each woman, according to need, spent in the delivery suite (from admission in labour to moving with her baby into postnatal care). Increased ratios of midwife time were allocated to women in categories III-V, these were respectively 1.2; 1.3 and 1.4.

BR+ grew from the application of the intrapartum score system in a number of maternity services (Ball and Washbrook 2010a). It did not however extend into recording midwife time in direct and indirect care beyond the intrapartum area. According to Ball and Washbrook (2010a:528) it had been proven to be robust and reliable. However, Sandall et al. (2011) recommended that the tool needed validation in supporting its future development and use.

BR+ was implemented by 101 maternity services across 117 sites in the UK in 2003 (Ball et al. 2003). Between 2003 and 2008, BR+ studies have been undertaken by 150 district general hospitals; tertiary maternity services; a number of small maternity services and stand alone midwife-led units in England (Ball and Washbrook 2010a). In addition studies were conducted in Wales, Scotland, Northern Ireland, and outside of NHS, in Eire and in New South Wales, Australia (Ball and Washbrook 2010a). These studies explored local staffing needs and budgets; and allowed for clinical workload reviews and changes in deployment of staff. It became apparent that although no precise information existed on how many establishments increased their staffing as a result of BR+, discussions revealed that many services either increased their overall staffing, or balanced shortages by redeployment of existing staff; by increasing their support staff or bank/agency staff (Ball and Washbrook 2010a).

It was recognised that the length of time (4-6 months) and cost of conducting a full retrospective BR+ study were a major limitation for managers of maternity services.

Workload management was also complicated by the dynamic work pattern in delivery suites (women give birth at any time – day and night and the unpredictability of emergency cases; thus workload can vary greatly from day to day or even hour by hour). Thus a new tool – the Birthrate Plus Acuity tool – was designed, which enabled midwives with a prospective and more predictive measure of women needs and staff required to meet those needs (Ball and Washbrook 2010b). It was first piloted in Wales (in 2006 & 2007) and consequently used in a variety of maternity services. An 'acuity' measure (measuring intensity of need within a delivery suite) was added to the original BR+ methodology to help managers to *'assess, compare and record fluctuating client workload with midwife availability in real time'* (Ball and Washbrook 2010b: 780). The tool incorporated a prospective classification of need; the number of midwives required; took into account the care provided to other women in the delivery suite and post-operative and postnatal care. A category of need was recorded by midwives on admission of each woman, which could be modified if needed during labour, at delivery or post-delivery emergencies. This allowed for the build up of 12-hour scenarios on the delivery suites, by recording all women and their needs in a 1, 2 or 4 – hourly intervals. The data were entered in excel spreadsheet which calculated the total delivery suite acuity and the number of midwives needed to meet it; it then compared that number with the midwives available. An analysis of this data overtime gives information on how often acuity and staffing are balanced or mismatched; thus highlighting the need for measures to be taken when demand outstrips supply. Analysing this real time data also could facilitate creating policies for minimum acuity/staffing levels that have to be maintained at all times – for example staffing to meet at least 85% of the recorded acuity (Ball and Washbrook 2010b).

Appendix II shows the midwife-to-woman and midwifery care assistant-to-midwife ratios in varied birth settings based on case mix categories and their definitions, adapted from BR+ (Ball 2006) by the RCOG 2007 *Safer Childbirth* report (p.29). The recommended ratios were 28 births to 1 WTE (whole time equivalent) midwife in a hospital setting, and 35:1 for home births. For birth centres/midwifery-led units,

the recommendation was for 35:1 (as these were normal deliveries of lower dependency women), not including transfers in cases of emergencies. Regarding caseload midwifery practice, where midwives provide the total antenatal, intrapartum and postnatal care, the recommendation was for 35:1. In labour wards, the recommendation was for 1.0-1.4 midwives for each established labour depending on case mix.

The headcount numbers of midwives differ depending on the sources used. The Nursing and Midwifery Council (NMC) keeps a record of all registered midwives, including midwives employed by local authorities, independent midwives and agency staff not employed by NHS, midwives in higher education and registered but not practicing midwives. According to the NMC there were around 30,000 registered midwives in 2008 (30,924 in 2010), while the number submitting intentions to practice (ITP) in England in 2008 was 28,030 (RCM 2010a). These numbers do not reflect midwives actually practicing. CfWI in 2012 based on information from the NHS Information Centre reported there were 26,825 midwives, a FTE of 20,790 in England as of April 2010. The NHS IC census captures the midwifery workforce employed by the NHS only (CfWI 2012), thus the smaller HC numbers reported. RCM submitted a memorandum (HS41) to Parliament in 2011, establishing that the country needed the equivalent of 4700 extra FTE midwives (CfWI 2011). This was based on *Birthrate Plus* workforce planning tool. The estimates of CfWI in 2011 were for additional 3,556 FTE of midwives.

There was a 16.4% increase in the midwifery workforce between 2003-2013 in headcounts and 19.2% in FTE (HSCIC 2014). The September 2013 figures showed that there were the equivalent of 1158 more full-time midwives in the NHS in England in that month, compared to September 2010 (HSCIC 2014). The RCM message over the last few years on need for investment in maternity services and generally in more midwives might have been partially acknowledged by the decision-makers.

2.7.1.2 AGING AND RETIREMENT

It is widely recognised that the UK midwifery workforce is ageing. The average age of a midwife in 2010 was 44 years (NHS IC 2010). In addition two thirds of the workforce was over 40 and a quarter was over 50 years, with fewer younger staff to replace them. The NMC estimated in 2010 that in the next ten years 40-45% of the midwifery workforce was going to reach retirement age (CfWI 2012). Future training of the workforce was also seen problematic as the majority of midwives who train were also approaching retirement age. The ageing and retirement trend is expected to have major implications on the supply of midwives unless there were sufficient numbers in training to compensate. However the actual number of commissioned places in 2010/11 was lower than planned (CfWI 2012).

2.7.1.3 WORKING TIME

As the UK midwifery workforce is female dominated it is perhaps not surprising that part-time working has increased over the ten years between 2000 and 2010. The number of midwives working full-time fell from 52.9% to 43.4% in the period 2000-2010 (CfWI 2012). Part-time work allows a younger workforce to combine work with family care. This may also create additional pressure on the remaining staff in understaffed units. It was self-reported that 85% of midwives had additional duties (due to vacancies, recruitment issues) which increased their workload and two thirds reported having to work longer than the agreed contracted hours (RCM 2010a).

2.7.1.4 RETENTION AND VACANCIES

According to the Midwifery 2020 report, midwives working for more than 5 years were more likely to stay in the profession and then retire. There was a high rate of midwives leaving their posts within the first five years of practice. The number of return-to-practice (RTP) midwives increased in 2009/10 (CfWI 2012); they are less expensive to employ than to train new midwives, even though their salaries are higher than for newly qualified midwives. Retention of experienced midwives remains a significant issue.

London Strategic Health Authority (SHA) had the highest number of midwife vacancies in 2010 (nearly 150) compared to other SHAs, but as a proportion of its workforce this number was low. According to a NHS IC survey[48] in 2010, the number of all midwife vacancies in England was 2.7% (559), which was slightly higher than in 2008 2.1% (407). RCM[49] reported fewer vacancies in 2010, for example 38 for London SHA (CfWI 2012). NHS IC definition of a vacancy is if a position has been vacant for 3 months. RCM on the other hand defines a position as vacant if a need has been reported by Head of Midwifery.

2.7.1.5 CHANGES IN EDUCATION AND TRAINING

The funding for clinician training is provided by Department of Health through a Multi-Professional Education and Training (MPET) levy (around £4.9bn in 2011/12, CfWI 2012). The DoH is in charge of the key priorities, while Strategic Health Authorities (SHAs) were in charge of planning and budget allocation based on local workforce needs. Consequently, providers were paid by SHAs which commissioned education, training and clinical placement. SHAs were accountable to DoH (via a Service Level Agreement) (CfWI 2012).

Practicing midwives need to be registered with the Nursing and Midwifery Council (NMC 2002)[50], after acquiring their midwifery degree and achieving all competencies required by the NMC. Previously qualified nurses are allowed to complete a pre-registration midwifery short programme in 78 weeks. There has been an increase in commissioned training places for midwives since 2002/3, with preference for degree qualifications compared to diplomas (CfWI 2012). Midwifery

[48] NHS IC (2011) Vacancies Survey March 2010 (online).

[49] RCM Evidence to the Pay Review Body.

[50] NMC replaced UKCC (United Kingdom Central Council for Nursing, Midwifery and Health Visiting established in 1983) in 2002. NMC maintains a register of UK nurses, midwives and health visitors, provides guidance and quality assurance and handles professional misconduct complaints.

pre-registration qualification is all currently at degree level (degree level education in nursing was introduced in 2013 for all students). There are three-year and four-year degree programmes for midwives. According to the RCM, the attrition rate is high – around 20% do not finish their studies and another 5-10% quit in the first 18 months of practice (CfWI 2012). Government policies support the need to increase the number of midwives, though there is no clear strategy of how to achieve it. The *Front Line Care* (2010) report, which focused on achieving high quality compassionate care, recommended to the *Midwifery 2020* panel to consider whether all midwives in leadership and specialist practice roles need to achieve a relevant degree; and that all midwives should be supported if they wish to obtain degree. Revalidation and increased investment in professional development were also suggested.

2.7.1.6 STAFFING IN MATERNITY UNITS

In February 2008, the Secretary of State for Health announced a package to support the recruitment of extra 4,000 midwives by September 2012 (NHS WRT 2009). *Towards Better Births* (HCC 2008) report reinforced the need for increasing the number of midwives.

Staffing in Maternity Units report was commissioned by the King's Fund in 2011. It was in response to concerns about safety and levels of staff in maternity services and aimed to explore whether improvement in safety could be achieved through better deployment of the existing staff, but recognised that just increasing staffing numbers may not lead to improved safety. Safety of care for mother and baby only in the intrapartum period were explored. The key findings of the report were that staffing mix, experience and deployment were significant factors and more important than staffing numbers alone, particularly in *"out of hours"* provision of care; and potential for certain task-shifting (routine examination of the newborn) between medical staff and midwives, without compromising safety and quality were identified. Nurses and neonatal nurses and maternity support workers (widely used across maternity services, with variation in level of training provided (Stout 2007) were identified as

potentially releasing midwife time. Midwife-led care was found to offer better outcomes for low and medium risk women; had a potential for cost saving and a suggestion for this model of care to be deployed across the service (in order to free obstetricians to care for women with complex need) was made. Continuity of obstetric- and midwife-led care (particularly caseload midwifery) was positively associated with safety and quality of care. Mode and place of birth as well as length of stay were expected to have implications for staffing requirements.

2.7.2 JUNIOR OBSTETRIC STAFF

Two factors have raised concerns about the experience and competence of junior obstetric staff. These are a shortening in the training and restructuring of career paths for new doctors (*Modernising Medical Careers* 2005), and the effect on training of the reduction in working hours for trainee doctors due to the European working hours regulations (the *Temple Report* 2010).

The European Working Time Directive (EWTD) was introduced in 1998 and fully implemented in UK in 2009 to include doctors in training. It limits work to 48 hours per week averaged over 6 months.

The RCOG report *Safer Childbirth* (2007) outlined the junior obstetric staff roles and the provision of more consultants to provide training and supervision to trainees due to the changes in the career paths for new doctors. The report found that in maternity services, junior medical staff practiced shift working, where effective handover arrangements were seen as a requirement for continuity of care and improved safety. The lead consultant obstetrician on the labour ward is responsible for planning efficient skill mix composition. The old system of junior staffing grades was divided into Senior House Officers (including career trainees and GP trainees) and Specialist Registrars (divided into years 1-3 and 4-5). Modernising Medical Careers (MMC) was introduced in 2007, which restructured the postgraduate medical training. The new programme is competency-based rather than time-based. It starts with a 2 year foundation programme (Foundation Year Trainees), followed by acquiring a place in

a specialist training programme (competition based), which continues for a minimum of 7 years: Specialty Trainees - ST1-ST2 (basic); ST3-ST5 (intermediate) and ST6-ST7 (advanced). Trainees enter the specialty training with little or no experience; during the basic training (ST1-ST2) they require intensive supervision and training (from midwives, senior trainees and consultant obstetricians) and are not in a position to provide service. Once competent at basic level they move to ST3-ST5 (middle grade rota) and acquire additional competencies but with less supervision. They are expected to have enough experience (visible from their individual logbooks, reviewed by their trainer with each attachment to a unit) to make basic decisions on the labour ward and perform some operative births but not to train ST1-ST2 trainees. The middle grade rota, given the limited working hours and lack of experience, will still require substantial daytime training and supervision as well as night-time supervision from consultant obstetricians. Unless provisions were made for extra consultants, the labour wards would be exposed to extra risks and the consultants' workload increased. Senior trainees (ST6-ST7) are expected to perform most operative procedures and routine decision making on the labour ward, with little supervision and increased responsibilities. Acquiring special and advanced labour ward skills and competencies in addition to training would be needed in order for them to take a specific consultant post in obstetrics and gynaecology.

A RCOG report, *The Future Role of the Consultant* (RCOG 2005), recommended that smaller units with less than 2,500 births a year should have 1 consultant and 1 specialist trainee (with at least 12 months experience in obstetrics and gynaecology). In larger units with more than 2,500 births a year, the consultant obstetrician should be supported by 2 or 3 specialist trainees. In larger units, the requirement for junior staff presence would be determined by workload and training opportunities and based on increased responsibility of these units to provide basic and advanced training in high risk obstetric practices. The regional training requirements will inform the need of the trainee's seniority. Protected consultant time for supervision was also required.

It was recognised that regional variations in the distribution of medical workforce existed and the future allocation of junior doctor posts should address this inequality in provision (NHS WRT 2009).

A general concern existed regarding the quality of training for all staff and the provision of safe care. Traditionally England had followed an *"experiential"* medical model of training, which was expressed in trainees spending long hours in their work place, during which time they acquired skills and knowledge. The EWTD has reduced these hours to 48 per week and this presented a challenge to deliver high quality training within the current service context. Accepting that majority of training should take place in a service environment; a change to the "*experiential"* model was needed so that NHS could continue to produce well-trained and safe professionals. The Temple report *Time for Training* (MEE 2010) was seen as having implications on the organisation of medical workforce, continuity of safe patient care and provision of out of hours care. The review was led by Prof Sir J Temple and commissioned by the NHS Medical Education England (MEE 2010). The purpose was to establish the impact of EU working time directive (EWTD) on quality of training for doctors, dentists, healthcare scientists and pharmacists in England.

The report established that EWTD predominantly affected doctors working in hospitals; it had greater impact on specialties with high emergency and/or out of hours workloads; that high quality training could be delivered in 48 hours, but was questionable in situations where trainees had a major role in out-of-hours service; when supervision was poor and access to learning was limited. Out-of-hours gaps in rotas were seen as the reason for not enough training and supervision of trainees; trainees were moved to cover night shifts at very short notice thus interrupting their daytime scheduled training. The gaps in rotas were affected by the introduction of full shift working patterns, which required more doctors; the full shift pattern of work also allowed very little trainer:trainee interaction and more handovers. In addition emergency care workload had increased and impacted negatively on the training opportunities for all trainees. Higher trainees were expected to cover in emergency rotas and unable to attend specialist training. It was also recognised that

the 15,000 hours available to trainees working a 48-hour week in a seven-year training programme, were not utilised effectively for training; the expansion in traditionally trained junior doctors and trainees in recent years working in traditional ways was not matched by quality and quantity of training opportunities; and training opportunities were even more negatively affected by the introduction of the New Deal. The role of the consultants in relation to training and supervision of trainees was also questioned. Out-of-hours services were still mainly delivered by trainees, with limited supervision, despite the expansion of consultants' numbers and the evidence of efficiency savings and improved patients' safety from increasing their presence. Training was not prioritised by consultants, for which traditional training models were better suited. A long list of recommendations followed which included: increased consultant presence 24/7 – increased responsibilities for direct care as well as increased flexibility to deliver training and supervision including out-of-hours; clinically responsible doctors to be employed in posts under a consultant contract, instead of expanding the grades; mentoring and support of trainees and consultants; training to be delivered in a reconfigured service environment under a consultant; reconfigured/redesigned elective and emergency services and effective Hospital at Night programme were seen as ways of changing healthcare to support training and safe services; the use of multidisciplinary teams to provide training; rewarding trainers; and training outcomes to be included as part of quality assessment of provider institutions (MEE 2010).

In the few cases where EWTD had been effectively implemented there was evidence for a positive impact of reduced working hours, mainly in relation to better organised rotas and work-life balance; more involvement of experienced doctors in acute care situations; effectively organised training in out-of-hours had an impact on patient safety and reduced loss of daytime training (MEE 2010).

To compensate for reduction in trainee numbers and hours and a rise in case mix complexities there was a call for 24 hours consultants cover[51] on labour wards (The *Future Role of the Consultant,* RCOG 2005; The *Future Workforce in Obstetrics and Gynaecology,* RCOG 2009).

2.7.3 OBSTETRICIANS

2.7.3.1 24 HOURS PRESENCE OF CONSULTANT-OBSTETRICIAN ON THE LABOUR WARD

The role of the consultant obstetrician is to guarantee provision of safe, high standard care for women and babies with complex medical or obstetric needs and to be present to respond to emergencies in an unpredictable environment (RCOG 2007). The implication is that greater presence within 24 hours is expected to lead to better outcomes, better decision making and a safer environment for birth. In addition the responsibilities of consultants extend to training, support and supervision of trainees; facilitating effective teamwork in multidisciplinary teams; involvement in risk management, compliance with clinical standards, diagnostic expertise and experience. The duties involve labour ward rounds, perform procedures (caesarean section), and respond to referrals from midwives. The expectation is that at a minimum there should be twice daily ward rounds on Saturdays, Sundays, bank holidays and one round later in the evenings. In emergencies if not physically present, the consultant should be contactable and must attend.

The report *Safer Childbirth* (RCOG 2007) revealed that only 30% of the maternity units claiming 40 hours consultant cover on the labour ward, actually had a consultant presence for that amount of time. Evidence from the RCOG *Hospital Recognition Committee* (2007) showed that only about 50% of maternity units in UK with births between 2500-4000 a year had a 40 hours consultant cover. 40 hours

[51] *Safer Childbirth (2007)* recommended 40 to 60 (for units with more than 5000 births per annum) hours per week presence of obstetricians on labour wards.

consultant cover was first recommended in *Towards Safer Childbirth* in 1999 and consequently adopted as a standard by the CNST.

In 2007 *Safer Childbirth* (RCOG 2007) replaced the term *"consultant cover"* with *"consultant presence"* and recommended that maternity units with more than 6000 births a year were to provide 60 hours of consultant presence; units with 2500-6000 births or classed high risk were to provide minimum of 40 hours of consultant presence; units with up to 2500 births a year were strongly advised to have 40 hours consultant presence but were allowed to base that on individual risk assessment. These recommendations followed a previous report *The Future Role of the Consultant* (RCOG 2005). It proposed obstetric staffing targets for consultant presence of 168 hours a week for units with more than 6000 births to be implemented by 2008 and for units with 5000 to 6000 births by 2010. This level of cover required 2700 FTE consultants by 2010. The NHS IC Census in 2008 recorded 1570 headcount and 1492 FTE consultants in obstetrics and gynaecology (O&G), (NHS WRT 2009). The gender workforce composition of consultants in O&G was 33% female and 67% male and as the age profile indicated that 50% of the consultants in this specialty were 40-49 years old, immediate retirement concerns were not expressed (NHS Workforce Review Team, 2009). There was a concern though about a potential increase in part-time working in the future due to an increase in numbers of female obstetricians, workforce trained under WTD and generational differences in attitude to work (work-life balance).

It was accepted that number of births did not reflect the number of complex cases which required consultant involvement; a scenario of reduced normal births within the obstetric units due to creation of maternity networks, may still leave the obstetric units with the same number of complex cases and the same demands on obstetric involvement.

The need for an increased consultant presence on the labour ward was driven by the increased number of obstetric interventions, evidence for increased infant mortality at night (NPSA, 2006 suggested that fetal compromise was more likely between

20.00 to 04.00 hours), the rise in number of births and complex pregnancies and the necessity to respond to emergencies as the activities on a labour ward do not vary much within 24 hours. These issues were amplified by a cohort of less experienced middle-grade staff and demands on consultants time to train and supervise (RCOG 2007). It was also recognised that though the rate of interventions[52] have increased there was no evidence for improved maternal or infant outcomes (RCOG 2007). The rise in intervention rates may have been influenced not just by the rise in complex cases but also by the relative lack of experience of doctors in training who provided the majority of obstetric care on the labour ward. There was evidence in the literature on the importance of the operator's experience in operative births, particularly in relation to maternal and neonatal morbidity linked to operative birth in the second stage of labour (Murphy et al. 2003).

Having a 24 hour consultant presence on the labour ward was seen to impact on job plans (this has to be included in the obstetricians' agreed job plan), remuneration, facilities, time-off allocation and work-life balance. The cost-effectiveness was also questioned. It was not conclusive whether the continuous presence would improve outcomes to the point where this presence was cost effective (RCOG 2007).

2.7.4 MATTERNITY SUPPORT WORKERS

The call for regulating support workers was made in *Front Line Care* report (Prime Minister's Commission) in 2010 in order to ensure high quality safe care. The role of maternity support workers is seen as complimenting and supporting midwives but not in substituting them (RCOG 2007) as they are not qualified to deliver babies. Their development, training, job titles, range of activities, grades assigned and pay rates varied across UK (Stout 2007). NICE (2007) believed that the role and contributions of MSW were unevaluated in UK. Key reports were published by the

[52] Now more than a third of women have either caesarean or instrumental births. The CS rate had increased from 10.4% in 1985 to 24% in 2010.

NHS National Workforce Projects (2009) and by the RCM (2010b) acknowledging their important contribution and suggesting wider use should be made of them. There was evidence from a Cochrane review (Hodnett et al. 2013) that support in labour provided by unqualified women may reduce interventions and improve maternal and neonatal outcomes. There was little evidence to justify the use of maternity support workers in provision of postnatal home visits. Morrell et al. (2000) found no health benefit at six weeks and six-months follow up and in use of the NHS service; the additional cost of a support worker in addition to routine midwifery care was £180 per woman.

The RCM estimated that the ratio of midwife to support worker was 5:1 (CfWI 2012). A skill mix optimisation and greater involvement of MSW in midwifery-led care and maternity care were seen as possible (CfWI 2012). More studies were needed to evaluate the impact of support workers on outcomes for mothers and babies; cost-effectiveness of the role at different levels of training and the views and experiences of women receiving care from them (Sandall et al. 2011, King's Fund).

2.8 CONCLUSIONS

This Chapter presented the complex interactions of policies, demographic, economic and organisational issues impacting on the maternity workforce.

Currently there are a number of major issues impacting on the maternity workforce. These include: rising birth rates; increasing complexity of births related to factors such as obesity and older first-time mothers; increase of interventions and specifically caesarean section rates; high obstetric litigations; aging workforce; changes to the service delivery models of care (with a move towards midwifery-led units) and patient centeredness with the choice guarantee agenda. At the same time there is a wide spread recognition of staff shortages (midwives and consultants) and doubts regarding provision of high quality, safe care.

Recognising the complexity of the issues affecting the maternity workforce, the House of Commons Health Committee in 2007 called for an integrated and cross-professional approach to workforce planning (House of Commons Health Committee 2007), aiming at having the *"right people, with the right skills, in the right place, at the right time"* (Workforce Review Team 2009).

However the latest House of Commons Public Accounts Committee (2014) report acknowledged that despite the published strategy for maternity services seven years ago (*Maternity Matters* 2007), Department of Health "*still has little grip in key areas and little assurance about performance"* (p.3). The Committee revealed confusion around the Department's maternity policy and described one of the main policies regarding continuity of midwifery care more as an aspiration than a clear objective. In addition neither the DoH nor NHS England was able to state who was responsible for ensuring NHS had enough midwives.

Given the economic constraints and the complexity of the above described issues it is of interest to understand whether staffing levels impact on the quality of care, and if so, how staffing could improve outcomes. The next Chapter will present the empirical evidence on the relationship between staffing and outcomes in maternity care.

3 CHAPTER 3 MATERNITY CARE OUTCOMES

This Chapter critically analyses the literature on the relationship between maternity staffing and maternal obstetric outcomes.

3.1 OBJECTIVES OF THE CHAPTER

As demonstrated in the previous chapter, staffing is an important policy issue in UK maternity healthcare because of perceived shortages impacting on the quality of care, coupled with the desire to minimise costs and offer a better choice to women.

This Chapter reviews critically the existing empirical evidence on the relationship between staffing and outcomes in maternity care, to determine whether staffing levels impact on the quality of care, and if so, how staffing could improve outcomes.

The Chapter thus identifies outcomes potentially sensitive to the quantity and quality of maternity care, which informs the selection of variables for a suitable model in the analytical part of this thesis.

The main focus was to identify outcomes related to the intrapartum period, which also impact on women's experiences of birth. This Chapter also addresses the differences between structure, process and outcome indicators and examines the relationship between these three, to determine how they are linked to quality of care.

The main questions addressed in this Chapter are:

- What is the difference between an outcome and a quality indicator in healthcare?
- What constitutes a maternity outcome and/ or indicator?
- What is the evidence for the relationship, if any, between levels of medical and non-medical maternity staffing and outcomes in maternity services?

3.2 BACKGROUND

Various studies have explored the effects of models of midwifery care on maternal outcomes in the UK and other countries with similar midwifery models of care (Sandall et al. 2013); the effects of continuous support during childbirth on outcomes (Hodnett et al. 2013); issues of safety of maternity services (Smith et al. 2009, King's Fund); place of birth (*Birthplace* 2011); staffing and outcomes in neonatal care (Redshaw et al. 1995); staffing in postnatal units (Yelland et al. 2006) and midwives' risk perception and intrapartum intervention rates (Mead et al. 2004).

It is of research and policy interest to identify the aspects of maternity care with the greatest impact on specific maternal outcomes, including organisational factors such as the types of maternity unit, and test whether maternity staffing levels help explain that impact.

However only exploring associations between maternity staffing levels and outcomes can provide a limited understanding if complex staff interactions are ignored (Gerova et al. 2010). Some of these relate to deployment of staff within organisations providing maternity care (for example between antenatal, labour suites, postnatal wards and in community setting) as well as skill mix. In order to maximise clinical and cost effectiveness (i.e. achieving the best outcomes for mothers and babies with minimum costs), it is essential to evaluate whether it is better to have more, less qualified staff or fewer more skilled and experienced staff and how best to organise them (Gerova et al. 2010).

A major question arising from the nurse staffing literature and outcomes is whether an association reflects a causal relationship. The U.S. Agency for Healthcare Research and Quality[53] (ARHQ 2007, Chapter 4:91) suggested that when higher

[53] U.S. Department of Health and Human Services.

staffing levels generate stronger effects for nurse sensitive outcomes compared to other outcomes, this could be a good indication (test) of a causal relationship.

3.3 NON-MATERNITY STAFFING LITERATURE

International research on the association between nurse staffing levels/skill mix and patient outcomes has been driven by recognition of global shortages of nurses and demographic changes affecting population (Buchan 2002). The shortages and the search for optimum staffing levels have been explored in parallel to other concerns such as patient safety and minimising risk (Rafferty et al. 2007; Torangeau et al. 2006; Naish 2006); in redefining the professional role of nurses in society and finding ways to measure the nursing impact (Griffiths et al. 2008; Maben and Griffiths 2008); and in relation to quality of care (Aiken et al. 2002).

Complexity of concepts such as "appropriate skill mix" and "minimum staffing levels" were also acknowledged in the literature particularly when considered in the wider economic, political, professional, socio-cultural or service users' contexts (Flynn and McKeown 2009). For example these two concepts may mean different things to health economists, managers or nurses depending on whether the goal was patient safety at all time, establishing a high ratio of registered nurses to healthcare assistants or making sure that no harm is done to patients (Flynn and McKeown 2009) therefore producing tensions within different contexts.

The majority of international literature on nurse staffing and outcomes has focused on acute service sector. Many studies have shown that low levels of registered nurses were associated with worse care outcomes, with a negative impact on the nursing workforce as well (Needleman 2002; Aiken et al. 2002; Rafferty et al. 2006; AHRQ 2006; AHRQ 2007; Lankshear et al. 2005). There has been a continuous attempt to develop and evaluate different models of nursing care by trying to link them to patient acuity or dependency, but none of the nurse staffing or skill mix models have managed to incorporate all the variables impacting on nursing workload or have shown a causal relationship with patient, nurse or organisational outcomes (Flynn

and McKeown 2009). It also seemed that most of the studies have explored the impact of 'poor' staffing levels or skill mix, rather than develop and evaluate new models of working (Flynn and McKeown 2009). The emphasis seemed to have been also on adverse patient outcomes in relation to nurse staffing levels and skill mix. These included in-hospital mortality, failure to rescue, pressure ulcers, medication errors, hospital acquired infections, post-operative complications, length of hospital stay (Rafferty et al. 2006; Garrett 2008; Tourangeau et al. 2007; Lankshear et al. 2005 and Kane et al. 2007).

Two systematic reviews (Lankshear et al. 2005 and Kane et al. 2007) aimed to assess the literature on the relationship between nursing workforce (levels and skill mix) and patient outcomes. Lankshear et al. (2005) synthesized the findings from 22 major studies mainly from the acute care sector between 1990 and 2004. All of the studies found an association with at least some of the outcomes investigated. Most of them strongly suggested that higher levels of nurse staffing and registered nurses dominating the skill mix were associated with improved patient outcomes (lower mortality; lower failure to rescue; less adverse events–falls, medication errors; reduced length of stay and reduction in complications-pneumonia, wound infections, urinary tract infections, etc.). It was not possible however to get a reliable estimate of the effect size of the associations and further research on the actual mechanisms by which nursing care affects patients was recommended. Some of the studies also emphasised that the quality of nursing care and better outcomes were associated with having more better qualified nurses and a higher proportion of registered nurses in the total staff mix. There were though indications of curvilinear relationship representing diminishing marginal returns to increased levels of registered nurses and skill mix, possibly meaning that with higher levels of better qualified nursing staff they may be increasingly involved with jobs that could equally be performed by other less qualified staff.

Most of the reviewed studies were cross-sectional (Silber et al. 1995; Aiken et al. 2002 and 2003; Tourangeau et al. 2002; Needleman et al. 2002; Jarman et al. 1999; Blegen and Vaughn 1998); and two were longitudinal (Mark et al. 2004; Unruh

2003). There were no studies based on randomized trials or using quasi-experimental methods, which in Lankshear et al. (2005) opinion if accompanied with insightful qualitative organisational research would have given almost an ideal platform to understand and evaluate how changes in nurse staffing levels impact on patient outcomes. The advantages of using experimental design were in obtaining robust results, i.e. less susceptible to confounding than observational studies. The next best option was longitudinal design, i.e. *"if the effect of nurse staffing is real, patient outcomes should change in relation to variations in nurse staffing over time"* (Lankshear et al. 2005). Attributing causality was also easier with longitudinal design as hospitals were seen as able to *"act as their own controls over time"* (Lankshear et al. 2005). Indeed the two longitudinal studies reviewed showed a correlation with increased registered nurses input over time and reductions in in-hospital mortality (Mark et al. 2004) and complications such as falls, pressure ulcers and urinary tract infections (Unruh 2003).

Some of the cross-sectional studies used large public administrative datasets and because of the representative nature of the patient level data used, the results were likely to be generalizable. However the data and methods used created limitations (i.e. robustness and causality). In cross-sectional studies, causality is an issue as other factors may also be influencing the results, for example whether hospitals with better outcomes attract better qualified nurses, or whether nurses prefer jobs in better performing hospitals. The reviewed studies varied in their case-mix and hospital characteristics adjustments. Only one study used multilevel modelling (McGillis et al. 2004) which takes into account the hierarchical nature of nurse staffing, hospital characteristics and patient data. Nurse staffing in most of the studies was considered in isolation from other staff groups particularly doctors. This omission was seen as important because hospitals with higher nurse staffing levels may have had more and better qualified doctors as well which may explain some of the variations in outcomes.

The studies that did consider doctors found an effect of nurse staffing to be over and above any association with medical staffing. The staffing measurements had their

own limitations: in some of the studies the staffing data did not distinguish between registered nurses involved in direct care and those having other roles (management, etc.); the staffing data were also aggregated at hospital level, rather than units/wards from which patient level data was extracted, leading to possibly compromising the accuracy, introducing random errors in staffing measurements and a decrease in the estimated effects; the distribution of registered nurses roles could have been different across hospitals with similar staffing levels; the data reflected paid hours therefore overestimating actual productive hours.

Lankshear et al. (2005) recommended that the future studies should be large, longitudinal, and possibly experimental, moving away from correlational studies in which it was difficult to establish causality and in which it was not known how the process or mechanics of nursing care affected patient outcomes. Both outcomes and staffing data should be collected at the lowest unit of care, i.e. wards and multilevel modelling should be used because of the hierarchical nature of the data. Qualitative organisational research should support quantitative analyses by enhancing the understanding of the causal mechanisms.

Kane et al. (2007) systematic review and meta-analysis aimed to assess the research evidence of the association between registered nurses (nurse-to-patient ratios) and patient outcomes in acute care hospitals. The review included 28 studies (17 cohort, 7 cross-sectional and 4 case control studies), which reported RN-to-patient ratios as independent variables and adjusted odds ratios of patient outcomes as dependent variables. Several outcomes sensitive to nursing care were included: in-hospital mortality, failure to rescue, patient falls, pressure ulcers, cardiac arrest, shock, urinary tract infection, hospital-acquired pneumonia, etc. The examined studies defined RN-to-patient ratios in different ways: number of RN FTE per patient day/or per 1000 patient days/or per occupied bed; number of patients cared for by 1RN per shift. The results from the pooled data of the reviewed studies revealed statistically and clinically significant association between increased RN staffing (additional FTE/per patient day or additional 1 RN per patient day) and lower odds of in-hospital mortality (for surgical, medical patients and in intensive care units, ICU); lower odds

of cardiac arrest, hospital acquired pneumonia and respiratory failure in ICUs; lower risk of failure to rescue in surgical patients. However the meta-analysis found that there was a greater cost-benefit from increased nurse staffing for surgical patients than medical patients.

The analyses in the examined studies were performed at patient or hospital levels. Studies using patient level data established larger effects of nurse staffing on mortality compared to studies using hospital level data.

Increasing the nurse-to-patient ratios has been recommended as a means to improve patient safety (AHRQ 2006). However there was recognition that the causal pathway to safe patient care included a range of other structure and process factors, such as hospital commitment to high quality and implementation of evidence-based clinical practice (Kane et al. 2007). Given that the use of large randomized controlled trials to establish the causal association with nurse staffing was unlikely, the authors accepted that inferences were to be based on observational studies, though these had to incorporate many other relevant factors such as patient and hospital characteristics and quality of medical care. One test of causality in observational studies was when higher staffing levels consistently produce stronger effects for nurse sensitive outcomes than for more general outcomes. In Kane et al. (2007) the effect of additional nurse staffing on failure to rescue and cardiac arrest, for example was higher than for mortality. Though other adverse outcomes, presumed to be nurse-sensitive (patient falls, pressure ulcers) did not show consistent association with nurse staffing.

Given this evidence from the research literature for a relationship between nurse staffing and outcomes in acute care, it was reasonable to investigate the relationship between maternity staffing levels and selected outcomes.

3.4 MEASURING QUALITY OF MATERNITY CARE

This section is informed by a reading of academic papers on clinical outcomes and risk factors in maternity. The purpose was to identify evidence for outcomes specifically sensitive to the quality of the maternity care. Answering this question requires an understanding of how healthcare quality was defined and measured.

There has been an increased interest in the last decade in measuring quality and safety of healthcare provision, with systematic evidence collected. Quality of healthcare in Lord Darzi's report *High Quality Care for All: NHS Next Stage Review* was defined as "*clinically effective, personal and safe*" (DoH 2008). Safety was considered the first dimension of quality and simply meant '*do no harm to patients*' by providing a clean environment, and reducing drug errors and avoidable infections. The personal dimension of quality related to provision of compassion, dignity and respect to patients. Clinical effectiveness was assessed not only by clinical measures of mortality, complication and survival rates but through patients' self-reported experiences of care and their post-treatment well-being.

Obstetric outcomes in the literature are evaluated with respect to providers' effectiveness and quality of care; investigated in terms of risk and patient safety; and result from a process of care (i.e. interventions and treatment for specific conditions produce certain outcomes). They could be influenced by patient characteristics, co-morbidities, models of care and place of birth, and could vary because of the providers' different philosophies, organisational culture, staffing levels and skill mix.

The spectrum of what was viewed as an outcome measuring quality of maternity care in the literature identified was wide, ranging from maternal and neonatal mortality/severe morbidity; outcomes measuring maternal adverse events ($3^{rd}/4^{th}$ degree tears, haemorrhage; etc); *"optimal"* events (i.e. normalcy of childbirth) or *"soft outcome"* (i.e. women's satisfaction with care). More outcomes from the midwifery literature exist, depending on whether the interest was in antenatal, intrapartum, postnatal or community care as midwives in the UK for example are

involved in all of these areas of care for the woman and her baby. Examples of outcomes in this respect could be: attendance for a 12 week antenatal scan; provision of one to one care in labour; normal birth; length of in-patient stay; mothers breastfeeding at discharge; re-admissions of mother and baby; mothers experiences and satisfaction with care; etc. (*Midwifery 2020,* 2010).

Two wider questions emerged:

1. What is the definition of and distinction between a quality indicator and an outcome?
2. Which indicators/outcomes relate to quality of maternity care?

3.5 INDICATORS VS OUTCOMES

Indicators are quantitative measures which reflect quality of care, they are tools for detecting problems and an opportunity to improve care but are not a direct measure of quality, as quality is multidimensional (Mainz 2003). Defining an indicator as a measure of overall quality will require investigating the causal relationship between structure, process and outcome.

Variations in outcomes are determined by variations in patient characteristics, structure and process of care and by random variations (Silber et al. 1995). Consequently, the quality of healthcare, however good, is only one of the many factors that affect patient outcome. As a result, using outcomes as indicators of overall quality of care is tricky. However if aspects of care process and/or structure with the biggest effect (clinical impact) on specific patient outcomes were identified, then resources could be allocated in improving these aspects (Rosenblatt et al. 1997), keeping in mind that certain patient characteristics such as age, parity and ethnicity will not be amenable to interventions.

The question about the distinction between indicators and outcomes was important because in the literature certain interventions, such as instrumental delivery,

caesarean section and episiotomy could be and are frequently investigated as outcomes, while they could also be considered as process of care indicators (caesarean section for example could be a process initiated by an obstetrician but could also be an outcome of a previous intervention/process and could have a beneficial or detrimental (or both) impact on maternal and/or infant health (RCOG 2013).

Furthermore if these interventions were accepted as process of care indicators, how were they related to outcomes and which particular outcomes (healthy mother/baby; maternal/neonatal mortality; severe morbidity; an adverse event – PPH, $3^{rd}/4^{th}$ degree tear)? The argument becomes even more complex when in some studies an intervention/process indicator was treated as a risk factor for an adverse event/outcome (for instance induction of labour has been reported in the literature as a risk factor for postpartum haemorrhage (PPH) (Knight et al. 2009); in instrumental deliveries, episiotomy was a risk factor for $3^{rd}/4^{th}$ degree tear, while having a caesarean increased the risk of haemorrhaging (Keriakos et al. 2013).

In relation to the links between interventions and outcomes, Griffiths et al. (2008) in their report *State of the Art Metrics for Nursing*, argued that identifying effective interventions and their associated outcomes was important, but because specific interventions will produce specific outcomes, interventions were unlikely to become indicators of the overall quality of clinical nursing services as "*indicators must be able to broadly reflect quality"* (Griffiths et al. 2008:16).

3.5.1 STRUCTURE/PROCESS/OUTCOME

Donabedian (1966) defined a structure/process/outcome classic model of healthcare quality. Structure relates to the characteristics of a healthcare setting and can incorporate staffing levels, deployment and skill mix, as well as organisational features such as facilities, equipment, management and culture. Process relates to what is done to the patient or how practitioners interact with patients in terms of diagnosis and treatment. Outcomes measure the results or consequences of care or

lack of care received by the patient and patient satisfaction of care (Iezzoni 1997). Outcomes are used to demonstrate that particular patient goals have been identified and achieved (Maloney 1999). The three components could be used to inform how to improve care unless there was a causal relationship between them in which case they can be used to measure quality of care.

Process measures have the advantages of directly measuring care received by patients and therefore could be used to detect if provision of care was good or poor. However that will depend on whether this process can actually be measured and whether data are available. Processes could also be manipulated particularly if linked to financial incentives and external assessment of performance.

"In order for a process indicator to be valid, it must previously have been demonstrated to produce a better outcome. Similarly, using structural indicators for quality assessment is possible only if structural components have been shown to increase the likelihood of either a good outcome, or a process that has previously been shown to yield better outcomes" (Mainz 2003:525).

Evaluating the process of care by which the outcome was measured and assessing the relationship between process of care and outcome was seen as important for quality improvement (Maloney 1999).

Maloney (1999) summarized clinical, economic and humanistic outcomes measurements. Clinical outcomes related to mortality, morbidity and effects of treatments; economic outcomes assessed direct and indirect cost of care as well as the economic implications of consumer satisfaction; humanistic outcomes related to specific outcomes consequence of specific treatment in relation to patient's quality of life. Patient's self-assessment of their own health status was also added to the outcomes investigated across disciplines.

According to Maloney (1999), there was no precise translation of the word *"outcome"*, meaning that there was no consensus on the best approach to define or measure an outcome; the term was not standardized across professions and

organisations; consequently the term outcome as applied to healthcare meant different things to different people. In most cases though the definition of outcomes was interpreted in terms of results (i.e. changes in patient's health status, condition or function following an intervention); costs or use of resources; and the categorisation was also influenced by the perspective of the data users. General health, disease-specific measures and functional status measures were the tools used to measure outcomes. *"The ultimate goal of outcomes research is to provide insights that lead to greater efficiency and higher quality of care "*(Maloney 1999:9).

3.5.2 DESIRABLE CHARACTERISTICS OF INDICATORS

Mainz (2003) established that indicators measure the extent to which set targets are achieved. *"They are expressed as numbers, rates or averages that can provide the basis for clinicians, organizations and planners aiming to achieve improvement in care, and the processes by which patient care is provided. They can be measures of structure, process and outcome, either as generic measures relevant for all diseases, or disease specific measures* (clinical indicators) *that describe the quality of patient care related to specific diagnosis*" (p.529).

Pencheon (2008) clarified that *"indicators help us understand a system, compare it and improve it".*

Quality of care and performance could be measured with indicators, which should be timely (i.e. using up to date data). The data itself should be complete and accurate. An indicator should be evidence-based and scientifically robust (the strength of the evidence determines its scientific strength); valid (i.e. it measures what was intended to measure); reliable (it can be reproduced across time and contexts); sensitive to change (it identifies important and substantial variations in clinical processes and outcomes) and interpretable (Mainz 2003).

Similar conditions for the usefulness of an indicator were presented by Armesto et al. (OECD 2006). Two main requirements were identified: importance and scientific soundness. Importance was determined by:

- Importance to health – whether a gap between the actual and potential levels of health exists and whether this gap was addressed by the indicator, for example data on mortality/morbidity could be used if a gap between actual and potential levels were identified;
- Policy importance - whether policy makers and consumers were concerned with a health area that the indicator measures, for example data on cost associated with a condition covered by an indicator could be used if public spending was a concern;
- Susceptibility to being influenced by the health care system – whether the indicator reflects parts of the health care system independently of confounders like patient risk and whether it can measure success or failure of policy changes – for example literature can be used to show whether the changes in the health system can influence a certain indicator.

The scientific soundness was determined by:

- Face validity – whether the indicator had basic logical and clinical meaning based on its past usage in national or other quality reporting activities;
- Content validity – whether the indicator captures meaningful aspects of the quality of care, assessed through a literature review of studies exploring the indicator;
- Reliability – whether the results were stable across various populations and circumstances, assessed through review of the literature.

In addition measurement itself was identified as an issue (Mainz 2003). The quantitative nature of indicators helps when comparing providers or over time to measure quality, safety and efficiency with the assumption that numbers are objective and more credible. But these quantitative measures will not necessary

capture the whole complexity of what, for example, midwives or other maternity staff actually do. This is not just because of the confounding factors but also because of that part of nursing (midwifery or other health professional) care that is *"invisible"* in nature and thus not easy to quantify (Duke and Copp 1992). Mainz (2004) also warned that not all aspects of care were measurable and sometimes aspects of care which were cheaper and easier to measure were of least importance for improving quality. He underlined that quality indicators had to be accepted as just being indicators of the quality of care with certain limitations; therefore the challenges were to identify, develop and implement indicators that expose as much as possible of the true quality.

3.6 STRUCTURE/OUTCOMES STUDIES

This section reviews studies investigating how outcomes are related to maternity staffing levels and actual birth place. The term outcomes in this section relate to what the authors defined as outcomes in their studies.

3.6.1 STAFFING

The literature search identified limited empirical evidence of association between maternity staffing levels and maternal outcomes. No study was identified which explicitly justified the choice of outcomes in terms of their sensitivity to the quantity of maternity staffing.

One relevant, cross-sectional study directly investigating the association between maternity staff (consultant obstetrician and gynaecologist (O&G), junior medical staff in O&G and midwives numbers per 1000 deliveries a year) and birth *"outcomes"* (caesarean section rate (CS), instrumental vaginal delivery rate (IVD) and epidural for labour rate) considered 1994-96 data for all Thames region maternity units (Joyce et al. 2002). Though the data were not recent, the issues discussed in the paper were relevant to the current debate of rising rates of obstetric interventions. Overall the results from the study's multifactorial analysis suggested

that staffing levels appeared unrelated to either epidural or IVD rates. Variations in epidural and instrumental vaginal delivery rate between units were most significantly explained by socio-demographic factors. Variations in CS rates on the other hand were related to the levels of monitoring and the experience of the obstetric staff, and independent of the correlation between caesarean and epidural rates. The level of junior, but not consultant medical staff, was positively correlated with caesarean section rates. There was no association between midwifery staffing levels and caesarean section rates in the multifactorial analysis after adjusting for confounders, which included epidurals, parity, induction rate and other.

A study from the United States (Hall et al. 2009) conducted a retrospective, descriptive analysis of association between the nursing care process (amount of nursing care received by patients), patient characteristics (age, race, marital status, number of previous births, history of previous CS, augmented labour, weeks gestation, ICD-9 admission diagnosis and peak acuity, represented by labour level severity) and patient outcomes (duration of three stages of labour; labour complications; CS delivery; fetal distress; patient length of stay and cost of care) at patient and unit level.

The contribution of this research was in its method to measure nurse care input (process of care), by estimating an hourly quantity of nursing care (referred to as *Nursing Effort* received by each patient) effect on outcomes. The analysis used *Nursing Effort Model* scores (hourly quantity of nursing care received by each patient, calculated from documented nursing observations and interventions) and demographic, financial, case mix and other clinical data related to each of the 900 women who delivered at three hospitals (small (1300 deliveries), medium (2300) and large (4200)) in Utah over a two month period in 2006. Two models of *Nursing Effort* were tested. In the first model at unit level each patient received the same *Nursing Effort* score equal to the total number of nurses to the total number of patients for each hour in each unit. In the second model at patient level, the score was

equal to the sum of the "*appropriate fraction of each documented nurse's effort, according to patient load, associated with each patient*"[54]. Only the scores calculated during the duration of active labour (the time between 4cm dilation and delivery) were used in the analyses. The strongest relationship between nursing care scores and outcomes were registered with the second model (derived from the individual nurse-patient interactions) and were related to the duration outcomes (durations of active labour; of complete cervical dilation to delivery and pushing duration to birth), thus suggesting that provision of additional nurse resources at key stages of labour improved labour progression and outcomes. All three labour duration measures were longer at the small hospital, for high acuity patients, for women giving birth for the first time and after augmentation of labour. There were no associations between nursing care and cost. The small facility, patients with higher acuity, CS and length of stay contributed to the higher costs. The study did not consider the skill mix of the nurses; outcome variables were treated as independent of each other; the determinants of patients' outcomes were limited; and there was an assumption of complete and accurate nursing documentation entry. The final results confirmed the hypothesis that the relationship between quantified nursing care and outcomes was better supported by patient-level measurement rather than by unit-level measurements. The authors suggested that unit-level measurements were useful for workload planning, but they did not capture nurse/patient interactions, nor the variations in the nursing resources distribution due to individual patient acuities.

Ashcroft et al. (2003) looked at the relationship between *"latent failures"* in the labour wards of seven maternity units in the North West of England and midwifery staffing, deployment and training, using a prospective semi-structured observational design. This method was created by using guidelines from RCOG and RCM,

[54] For example: A patient received documented care by nurse A and nurse B during 1 hour. Nurse A recorded care for 4 patients during that same hour and nurse B documented care for 5 patients during the same 1 hour. Thus the patient received 1/4 of a unit of Nursing Effort from nurse A and 1/5 of a unit of Nursing Effort from nurse B. The total nursing effort for this particular patient was 9/20 or 0.45.

standards from Clinical Negligence Scheme for Trusts, analyses of cases from the Confidential Enquiry into Stillbirth and Deaths in Infancy and obstetric claims and maternal mortality cases. The total observations lasted 7 days (48-52 hours' total observations) in 2000; all midwives on the labour ward were observed over 2 to 8 hours; other visits included delivery rooms, operating theatres, around fetal monitors and pH apparatus etc. There was also a follow-up visit for one day the following year after the study ended, which showed little change. *"Latent failures"* were defined as *"accidents waiting to happen"*.

All units experienced shortages of midwives and poor skill mix during the observed period, where the most pronounced shortages were experienced by the larger inner city units (six units used bank staff to achieve minimum staffing levels). All units had met level-1 requirements of Clinical Negligence Scheme for Trusts. At the time of staff shortages the use of epidural and oxytocin practices continued while it was known that these practices needed increased midwifery supervision. Situations were observed where shortages and poor deployment of midwives resulted in risk situations for mother and baby. These included: an emergency CS procedure delayed because all midwives on duty were busy caring for other women (particularly in second stage of labour) and could not assist in the maternity theatre (emergency CS policies at the unit normally required 2 midwives in theatre); induction for a twin pregnancy was unsuccessful and an emergency CS was required (which required 3 midwives to assist in theatre); newly qualified midwives were left to care for normal births on busy wards, a complication arose (shoulder dystocia), followed by delayed response because of staff shortages, resulting in adverse event (birth asphyxia). It was also observed that the use of team midwifery[55] had led to experienced midwives being deployed in community teams while less experienced midwives were working on labour wards dealing with normal and high risk cases. Issues with opportunities

[55] Team midwifery was introduced following publication of the report of the Expert Maternity Group entitled, '*Changing Childbirth' (Department of Health 1993)*, which included among its recommendations that 75% of women in labour should be cared for by a known midwife.

for training and updating skills (emergency obstetric management and interpretation of cardiotocographs) were also identified. These were only provided during working hours as shortages of staff prevented participation in formal training sessions. One of the most disturbing observations was that *"near misses"* because of staff shortages occurred frequently (on average one in every 2.5-5 days, usually in units with most deliveries and complications) and stayed unreported. *"Near misses"* were defined as events which under different circumstances could become an accident.

A report commissioned by the DoH Policy Research Programme and produced by the National Nursing Research Unit (NNRU) in 2010 (Gerova et al. 2010) aimed to assess the feasibility of using routinely available data to measure the impact maternity staffing has on birth outcomes in maternity services at NHS trust level in England. Patient-level data at trust level was selected from *Admitted Patients HES* data from Dr Foster for the period April 2008 – March 2009. This included 144 trusts out of 150 which then provided maternity care in England and 615,042 women who had given birth in hospital. The study only explored *"readmission within 28 days"* as an outcome, defined as number of women being readmitted to any hospital within 28 days after discharge from the postnatal ward. A limited selection of predictors was available: age and ethnicity of mother; Carstairs deprivation index; Charlson co-morbidity index; delivery type; professional overseeing the birth ; number of admissions in the previous 12 months; pre-and post-birth length of stay. The staffing variables were selected from the *Maternity Matters Benchmarking* dataset (2008) and matched at trust level to the *Admitted Patients HES* data. The "*FTE/ birth ratio"* was defined as number of births per health professional FTE. After partially removing some of the women's and trusts' contributions to the variations in the re-admissions, the aim was to find out whether the staffing variables would explain some of the remaining variations.

There was a significant relationship between all staffing variables and readmissions ($p<0.001$). The results showed that higher numbers of full time equivalent (FTE) midwives per birth; a higher ratio of consultant obstetrician FTE to midwives FTE and a higher ratio of consultant midwives FTE to midwives were all associated with

a lower probability of readmission. A higher ratio of registered nurses FTE to midwives FTE was associated with a higher probability of readmission. The authors admitted that the relationships demonstrated were certainly plausible with better outcomes consistently associated with higher levels of more experienced and more highly qualified staff. However risk adjustment was limited[56] in the models used and a possibility remained that further risk adjustment might alter the relationships. Further the authors discussed the limitation of modelling associations between staffing levels and outcomes without a consideration of the complex interactions involved. For example it was unclear how maternity staff, including registered nurses were deployed within trusts (for example between delivery suites, postnatal wards, operating theatres/recovery area and community). The question whether it was better to have more, less qualified staff or fewer more skilled staff and how staff should be deployed to maximise clinical and cost effectiveness remained.

Overall research directly investigating the relationship between maternity staffing levels and maternal obstetric outcomes was limited with conflicting results about the effects of staff levels, experience and deployment on different outcomes. These were mainly retrospective or descriptive studies, with incomplete risk adjustments and fewer outcomes. Some confirmed the perception that lower staffing levels and/or staff experience were associated with adverse outcomes in terms of safety. However these studies were not able to provide estimates of the impact of changes to staffing or provide robust evidence to guide policy about staffing levels. There were also limitations with respect to application to NHS care of studies undertaken in other countries.

[56] Variables not included in the risk model: previous delivery type, parity, multiple pregnancies, multiple births, gestational age and co-morbidities, as the Admitted Patients HES version did not contain the maternity tail, where some of the clinical data was available.

3.6.2 PLACE OF BIRTH STUDIES

An increasing number of studies compared maternal and neonatal outcomes and safety of births planned at home (under midwifery care) and hospital settings (usually obstetric care). Observational studies from the UK (Chamberlain et al. 1999; Symon et al. 2009; Mori et al. 2008), United States (Janssen et al. 1994), the Netherlands (Wiegers 1998; de Jonge et al. 2009), Canada and British Columbia (Janssen et al. 2009; Janssen et al. 2002) and Sweden (Lindgren et al. 2008) generally reported that planned home births usually for women classed as low risk before the onset of labour appear to be as safe as hospital births, with less obstetric interventions for women, although neonatal outcomes have not been consistent.

Two Cochrane reviews compared births in obstetric and 'alongside' or 'co-located' midwifery units (Hodnett et al. 2010) and home vs hospital settings (Olsen et al. 1998). The first (Hodnett et al. 2010) considered 9 randomised controlled trials (10,684 women) and concluded that midwifery settings offered women a higher probability of spontaneous vaginal births, maternal satisfaction and less interventions. In addition no association between the setting and perinatal and maternal severe morbidity and mortality was observed, though the review lacked sufficient power to identify differences in rare adverse perinatal and maternal outcomes. The second review included only one randomised controlled trial with 11 women and found no differences in outcomes and safety between home and hospital settings. A meta-analysis (Olsen, 1997) based on six observational studies considered safety of home births and perinatal outcomes for low-risk women (24, 092) and their babies. Women planning births at home were less likely to experience interventions (in terms of induction, augmentation, instrumental delivery, CS, episiotomy) and their babies were less likely to have lower Apgar scores, while there were no differences in perinatal mortality.

A study by Wiegers (1996a) examined the differences in obstetric outcomes of planned home and hospital births for low risk women in Netherlands. Using the concept of optimality (adopted from Prechtl, 1980 and described in detail later in this

chapter) she concluded that the outcome of planned home births for low-risk women is "*at least as good as the outcome of planned hospital births for first time mothers, while for other mothers the outcome of planned home births is significantly better*" (Wiegers 1996a). Longer labour, the need for sedation and neonatal problems for first-time mothers were more prevalent in the planned hospital births, while parous women with planned hospital births experienced more perineal tears, PPH, delayed progress in labour and other interventions. An interesting finding was that the differences in planned home and hospital births were explained mainly by structural factors and midwifery practice specifics and not as much by mothers' characteristics, case mix or medical conditions. Another large retrospective study from Netherlands (de Jonge et al. 2009), using routine data for 500,000 women examined perinatal mortality and morbidity for low-risk women and found no difference in outcomes among planned home and hospital births.

This was confirmed by Canadian (Janssen et al. 2009) and Swedish (Lindgren et al. 2008) studies for low risk women, which presented evidence for low obstetric interventions for home births. A prospective cohort study by Janssen et al. (2002) from British Columbia showed similar neonatal results (in terms of mortality, Apgar scores, meconium aspiration syndrome) for home (attended by midwife) and hospital deliveries (attended by physicians). The detected differences related to mothers (of similar obstetric risk, after adjusting for maternal age, lone parent status, parity, income, use/no use of substances), i.e. procedures such as epidural analgesia, induction, augmentation, episiotomy were less likely for women giving birth at home (n=862) under a midwife, compared to women giving birth in hospital under a physician (n=743) or in hospital with a midwife (n=571). The last three studies however were underpowered to compare differences for serious adverse outcomes in both settings.

An anonymised matched cohort study from Scotland (Symon et al. 2009) compared clinical outcomes for 8676 women cared for by an independent midwife (1462) and women using NHS services (7214). Clinical outcomes such as normal birth, spontaneous onset of labour, use of analgesia, perineal trauma and breastfeeding

were significantly better for women giving birth under the care of independent midwife. Neonatal outcomes such as prematurity, low birth weight and rate of admission to neonatal intensive care were all higher for the NHS cohort. However high risk women (with pre-existing medical conditions, previous obstetric complications, breech presentations and twin pregnancies) cared for by an independent midwife were more likely to experience still birth and neonatal death (the incidence of still birth and neonatal death for low risk women was comparable to other studies of low risk births). Also a higher proportion of independent midwives assisted women who delivered at home (66%) compared to the NHS cases (0.4%).

The Birthplace Research Programme[57] in England, led by researchers from the National Perinatal Epidemiology Unit (NPEU) at the University of Oxford commenced in 2007. The purpose was to fill a knowledge gap about the configuration and organisation of maternity services; safety, models of care and cost-effectiveness of different settings for birth[58] in England. This evaluation was driven by the policy direction of offering women more choice, including for place of birth (DoH 2004) and the need for high quality research comparing clinical outcomes and safety of different birth settings (NICE 2007). Results of outcomes for babies and women at low-risk of complications before the onset of labour for births planned at home, in different types of midwifery units and in obstetric units were reported in 2011 (Birthplace 2011). This was a large scale, prospective cohort study of over 60,000 women and classified as healthy with low risk pregnancies[59], if before onset of labour they did not have any of the medical or obstetric risk factors presented in

---

[57] Commissioned by NIHR Service Delivery and Organisation Programme and the Department of Health Policy Research Programme, UK.

[58] Obstetric Unit (OU), Alongside Midwifery Unit (AMU), Freestanding Midwifery Unit (FMU), Home births.

[59] Women who had elective caesarean section or caesarean section before the onset of labour, presented in preterm labour (<37 weeks' gestation), had a multiple pregnancy, or who did not receive antenatal care were excluded, as well as still births before the onset of labour.

the NICE 2007 intrapartum care guideline. These factors were considered to increase the risk for woman and baby, and care in obstetric unit was seen as reducing the risk.

The results confirmed that giving birth in England was very safe. Furthermore, safety, low interventions and similar outcomes for babies were validated for 'low risk' women giving birth at midwifery units, with only half the rate of CS compared to the obstetric unit births. Women giving birth in all three non-obstetric settings had lower odds of interventions, such as augmentation, ventouse or forceps delivery, intrapartum caesarean section, episiotomy. The proportions of women with 'normal birth'[60] were: 58% for planned obstetric unit; 76% for alongside midwifery unit, 83% in a freestanding midwifery unit and 88% for planned home births. The adjusted[61] odds for 'normal birth' were also higher for the three non-obstetric settings. Other adverse maternal outcomes such as 3rd/4th degree perineal trauma, blood transfusion and admission to higher level care were lowest for planned births in freestanding midwifery units but overall did not show a consistent relationship with planned place of birth.

The research warned about the higher risks for the babies of first time mothers giving births at home and for the higher transfer rates to obstetric units of first time mothers giving birth at home (nearly 50%) or in a freestanding midwifery unit (over a third). Births outside obstetric units were also more cost effective particularly for multiparous women (as it combined safety and cost effectiveness). This cost saving option was not without a drawback though - given that the main cost contributors were fixed unit overheads and staffing, offering one-to-one care at a midwifery unit or at home, would require employing more midwives which would add to the overall costs (at the same time shortages of midwives have been observed).

---

[60] Birth without induction of labour, epidural or spinal analgesia, general anaesthesia, forceps or ventouse delivery, caesarean section, or episiotomy.

[61] Adjusted for maternal age, ethnic group, understanding of English, marital or partner status, body mass index in pregnancy, index of multiple deprivation score, parity and gestational age at birth.

The Birthplace authors suggested that the changes in the configuration of maternity services should take account not only the costs but the issues of safety, transfers, occupancy rates, staffing levels (including support staff), skills, deployment and training. Overall one of the recommendations was that healthy women with uncomplicated pregnancies could safely be offered a choice of birth place. Apart from the scale of the study, the other valuable contribution was in the area of collecting additional information which was not routinely available (for example, percentage of NHS trusts in England with midwifery units in 2010); the configuration of the maternity services and models of care in England; the characteristics of organisations delivering high quality care (through individual organisational case studies); additional patient level data on interventions, complications, outcomes and unit-level data on staffing levels and skill mix. A major criticism remained however, that despite the scale of the study and the large cohort, the main focus was on low-risk women.

3.7 PROCESS/OUTCOMES STUDIES

Process refers to how practitioners interact with patients (Donabedian 1966), and the types of care, diagnosis and treatment that could result from a particular organisational structure, including staffing levels.

This section analyses process/outcomes literature in relation to quality of maternity care, including: comparison studies (midwifery vs obstetric care); midwifery models of care; continuity and one-to-one care; and interventions and outcomes studies.

The general view in the existing literature and among practitioners and patients is that midwifery models of care result in less medical intervention and invasive procedures compared to medical models of care. Studies suggested that women delivering under midwives were less likely to be induced or use oxytocin for augmentation, less likely to receive continuous electronic fetal monitoring, epidural anaesthesia or episiotomies; or to have an operative vaginal delivery (forceps or

vacuum), resulting in less perineal trauma (Chambliss et al. 1992; Rosenblatt et al. 1997; Sandall et al. 2013).

Maternal related outcomes were considered in the contexts of providers' effectiveness; in terms of risk and patient safety; by comparing outcomes for different providers (obstetricians vs midwives) and the roles and philosophies of different providers to interventions. Studies provided evidence that differences in the obstetric providers' outcomes were determined by their different philosophies, different approaches to intrapartum care, technical expertise and training. Midwives consider normal birth as their main philosophy (Rooks 1999). From a consumer perspective, midwives offered individualised care, support and encouraged family involvement, intimacy and normalcy of birth process. Large inter-specialty differences existed in the way providers treat similar conditions and even greater variation exists across disciplines (Rosenblatt et al. 1997).

Outcomes of midwifery care have been found not only to be as good as that of other obstetric providers but also cost effective (MacDorman and Singh 1998; Birthplace 2011). Midwifery care was less costly, used fewer resources and contributed to better patient satisfaction compared to physicians care (Hastings-Tolsma et al. 2009).

3.7.1 CONTINUITY OF CARE, ONE-TO-ONE MIDWIFERY CARE, MODELS OF CARE

Continuous support and one-to-one care from midwives are viewed as beneficial to women and are crucial to their positive pregnancy and labour experiences.

A review of maternity services in England (HCC 2008) recognized that a small minority of women (median 20%, range 9-34%) experienced care from the same midwife during labour or knew one or more of the staff caring for them (median 21%, range 10-43%). At the same time around 20% (range 11-40%) of women were left alone during labour which worried them. RCOG (2008) and NICE clinical guidance (2007) on intrapartum care recommended women in established labour and childbirth to receive one-to-one care by a designated midwife for 100% of the time.

With the adoption of *Changing Childbirth* policy in 1994 (DoH 1993), which emphasized the concepts of woman-centred maternity care, choice and control for women, continuity of carer, midwifery-led care and transfer of care to community; one-to-one midwifery care became one of the instruments for the implementation of this policy. In the previous decade, efforts to improve midwifery-led care and continuity of care, was through midwifery-led units and team midwifery (Wright et al. 1993). However evidence emerged of relatively high occupational stress and burn out in team midwifery (Sandall 1997). Furthermore Sandall (1998) suggested that the occupational burnout of midwives was related to having low control over decision making in their work pattern, longer hours shifts and lower grades. One-to-one midwifery models of care were introduced in 1993 and intended to compensate for some of the team midwifery shortcomings by providing more flexibility, control, autonomy and organisational support for midwives in their work as well as to facilitate meaningful relationships with women. In this model, a named midwife cares for each woman in their caseload through the entire antenatal, labour, birth and postnatal period, rather than being entirely ward or community based. The aim was the provision of woman and family-centred care through continuous and community evidence based midwifery care and through improvement of midwives' interpersonal skills.

A prospective comparative study by Page et al. (1999) evaluated the effects of this model in terms of continuity achieved, some clinical interventions/outcomes for women and babies, women's satisfaction of care, staff response to changes and its cost effectiveness. The study compared the existing traditional system (mostly consultant-led shared care) of care with one-to-one midwifery care in the Hammersmith Hospitals NHS Trust in London between August 1994 and August 1995. Randomisation of women and staff was not feasible, and the data were obtained through medical records (for 374 women in the study group and 528 in the control group), hospital statistics, interviews and questionnaires. Though women compared were of similar obstetric risk, the control group was more ethnically diverse and of lower socio-economic status. Higher levels of continuity of care and

women's satisfaction; lower levels of intervention (epidural, episiotomy) and better outcomes (fewer perineal tears) were observed for the one-to-one care group after controlling for maternal characteristics. Assisted delivery rates, CS and neonatal outcomes and breastfeeding rates were similar for the two groups. There was no increase in the rate of normal birth for the one-to-one care group and despite the lower rate of interventions for this group these remained high in relation to other maternity services. One-to-one midwives assisted in more births but this did not cost more than the conventional care. Overall one-to-one care was judged as safe, providing a higher degree of continuity and fewer interventions.

Many of these results were confirmed by a randomised controlled trial study of 1000 women in Australia (Biro et al. 2000) comparing continuity and outcomes of care provided by midwives (in team midwifery) with standard care (obstetricians and midwives). Continuity of midwifery care was associated with less intervention (augmentation, electronic fetal monitoring, epidurals and episiotomy) but more tears, and shorter length of stay. There were no differences in perinatal mortality. Other studies (Blake et al. 2001; Benjamin et al. 2001; Hodnett et al. 2013) confirmed the same positive outcomes for women (less intervention, including fewer epidurals, CS and instrumental deliveries) and an increased satisfaction of giving birth (Hodnett et al. 2013; Page et al. 2001) as a result of continuous, one-to-one midwifery support during labour.

Hodnett's et al. (2013) *Cochrane Review* of the effects of continuous support during labour included 22 randomised controlled trials, from 16 countries, involving 15,288 women. The results of the review revealed that women with continuous one-to-one support (provided either by nurses or/midwives; non-institutional staff such as lay people or people from woman's own support network), had a shorter duration of labour and were more likely to give birth without epidural analgesia or anaesthesia; to have a spontaneous vaginal birth; and less likely to: have instrumental or caesarean birth and be dissatisfied with their experiences. The included trials did not report any adverse effects – there were evident physiological and psychological

benefits for women giving births with one-to-one support without any accompanying risks.

In an earlier study Hodnett (1997) had looked at the reasons preventing midwives and nurses from providing effective one-to-one support benefiting women in labour. These were: the simultaneous process of looking after several women; managing technology and keeping records; working in understaffed wards and changing shifts in the middle of women's labours or lack of labour support skills.

The relevance for the research presented in this thesis is that the issues raised above are part of process and structure of care (including staff shortages), which impact on: midwifery workload and therefore the quality of care provided; on birth outcomes and experiences of women.

Evidence relevant to midwifery models of care includes a Cochrane review of 13 trials from Australia, Canada, New Zealand and the United Kingdom, which compared outcomes from midwife-led continuity models of care with other models of care for childbearing women and their infants (Sandall et al. 2013). Midwife-led care was associated with certain benefits for women with no identified adverse effects. The benefits included: less use of analgesia; fewer episiotomies or instrumental births; more spontaneous vaginal births; increased chance of being cared by a midwife they knew; being more in control during labour and; initiating breastfeeding.

Given the research evidence, clinical guidance and the positive experiences for women, one-to-one midwifery care in established labour and continuity of midwifery care were recommended as practices for increasing normal birth rates by Dodwell and Newburn (NCT 2010). Together with *Midwifery 2020* programme they suggested using normal birth as a quality indicator of midwifery care.

3.7.2 INTERSPECIALTY DIFFERENCES, PHILOSOPHY

Rosenblatt et al. (1997) looked at the interspecialty differences in the provision of prenatal and intrapartum care for low risk women. A random sample of low-risk patients was drawn from a random sample of obstetrician-gynaecologists, family physicians, and certified nurse-midwives between 1988 and 1989 Washington State (US). The results confirmed the initial hypothesis of high intervention practices among obstetricians, fewer obstetric interventions among nurse-midwives (less use of induction, augmentation, continuous electronic fetal monitoring, epidural, lower CS rates and fewer resources) and an in-between position for family physicians. The differences in practices, given the variations in the treatment of similar conditions, were attributed amongst other things to the economics and politics of medicine, litigation issues, women's preferences and healthcare professionals career choices, all of which affected healthcare professionals orientation, training, philosophy and skills which in turn determined process, management and quality of care delivered, outcomes, access, costs and women's experiences and satisfaction of care.

Some qualitative studies looked at how midwives' perceptions of intrapartum risk (Mead et al. 2006); healthcare professionals views of safety in maternity services (Smith et al. 2008); and midwives' experiences of midwifery practices in hospital environment (O'Connell and Downe 2009) impacted on the intensity of interventions. These studies suggested that there was a link (though weak) between high perception of risk and higher interventions and that midwives working in a higher intrapartum intervention environment had a higher perception of intrapartum risk and other way round. The study by Mead et al. (2006) observed that midwives had an over-pessimistic view of normal progression of labour and an over-optimistic risk perception of outcomes resulting from interventions. In terms of perception and barriers to safety in maternity, the study identified issues such as understaffing, inappropriate skill mix and poor morale, management, communication and training opportunities; and increased complexity of pregnancies. There was also recognition of the delicate position of midwives in between two belief systems – the

"biomedical" system of science and technology and therefore higher medicalisation of pregnancy and birth and the *"phenomenological"* - enhancing the woman's physical and emotional wellbeing. The so called *"hybrid"* midwife appeared to adapt to her work setting and changed her practice as a result of accumulated experience. Experience seemed to be quite important, as largely experienced midwives tried to normalise birth despite the setting or risk complexity. The study also suggested that the work setting and the institutional goals, values and resources prevented the provision of individualised care for women by midwives.

Brown and Grimes (1995) critically presented the general limitations in studies related to effectiveness of nursing and midwifery care when comparing outcomes of their care to those of care by physicians. These were divided into conceptual and methodological issues. Conceptual concerns included lack of sensitivity of outcome measure to detect expected changes in populations served by midwives and inattention to the relationships between the process of care and the outcomes of care. Methodological concerns included lack of randomized assignment of patients to providers in order to control for patient morbidity (the studies focused mainly on low risk women to control for heterogeneity); lack of information to enable assessment of internal and external validity; inadequate sources of data; imprecise standards of care and criteria for comparing patient outcomes; and superficial cost analysis.

3.8 OUTCOMES STUDIES – MORTALITY, MORBIDITY, OPTIMALITY

Process relates to how practitioners interact with patients in terms of diagnosis and treatment, as described in the previous section. Outcomes measure the effects of care on the patient and patient satisfaction of care, in maternity this may range from normal birth to serious morbidity or death. The perfect or ideal outcome indicator would capture the care process perfectly, without the interference of other influences on outcome such as patient characteristics.

As Mainz said, in his classic definition: *"An ideal outcome indicator would capture the effect of care process on the health and wellbeing of patients and populations"*

(Mainz 2003:525). Mainz (2003) defined outcomes as *"states of health"* or *"events that follow care"*, expressed as the *"five D's: death, disease, discomfort, disability and dissatisfaction"* (Lohr 1990, cited in Mainz 2003).

In the obstetric literature outcomes are frequently defined in terms of mortality and serious morbidity. This to an extent determined also reviewing the literature on maternal outcomes in terms of mortality and morbidity. In addition, the opposite spectrum, the literature based on the concept of optimality, is also presented.

3.8.1 MATERNAL MORTALITY

Maternal mortality is and has been a rare event, particularly among the healthy population in the high-income countries for the past few decades. The report of the triennial UK enquiries into maternal death led by CEMACH *Saving Mothers' Lives* revealed that the maternal death in the UK was extremely rare[62] and that in the triennium 2003-05 the maternal mortality rate (MMR) was 14 per 100,000 maternities. The main direct cause of maternal death was thrombosis and thromboembolism. The report also recognised that "*in only five of the 36 cases of women who had midwifery led antenatal care and who died of Direct, Indirect, Coincidental or Late Direct causes was midwifery care judged to be substandard*" and that there were "*no cases of poor midwifery care amongst the 16 women who died and who had received joint midwifery/GP led care (p.9)*". The main reasons mentioned for the substandard care by midwives, were poor diagnosis and management of medical conditions; failure to seek obstetric advice or being ignored by other medical staff if such advice was sought. The most recent report of the triennial enquiry (CMACE 2011) revealed that the MMR for 2006-08 in UK was 11.39 per 100 000 maternities[63]. The decline from the previous triennium was

[62] 295 women died from causes directly (n=132) and indirectly (n=163) associated with pregnancy (2003-05, UK).

[63] 261 women died from causes directly (n=107) and indirectly (n=154) associated with pregnancy (2006-08, UK).

attributed to reduction in death cases from thromboembolism and haemorrhage. This time though sepsis was the main direct cause and cardiac disease the main indirect cause of maternal death. More than half (63%) of the women identified by the enquiry have died in the postnatal period. It was observed that in many avoidable situations and cases midwives "*failed to carry out and act on basic observations*" (Garrod et al. 2011).

The low incidence of maternal mortality does not allow identification of significant changes over time as well as comparison between subgroups of women (Pattinson et al. 2005). Studies of mortality also need large populations; they are complex, expensive and need long-term perspective.

Maternal mortality is an exceptionally important event and subject to serious investigations. This outcome because of its rare incidence in the UK and for the reasons mentioned above was not considered in the current research.

The following paragraphs will explore the literature on quality of maternity care in relation to outcomes of maternal morbidity (namely adverse outcomes such as severe perineal tears - $3^{rd}/4^{th}$ degree tear and post-partum haemorrhage); and optimal outcomes (i.e. achieving best outcomes for mother and baby with least interventions). Optimality concept was presented as an alternative to morbidity - it measures process and outcomes of maternity care, from the position of health rather than illness.

3.8.2 MATERNAL MORBIDITY

Maternal morbidity has been investigated in previous empirical research, within the framework of adverse maternal outcomes or '*near misses*'. Severe maternal morbidity has been considered by researchers as a better indicator of quality of maternity care because of the rarity of maternal mortality in the developed world. A WHO systematic review (Say et al. 2004), identified the prevalence of severe acute maternal morbidity (i.e. '*near miss*' – intuitively defined by most studies as a woman

who almost died but survived) in the world to range from 0.01% to 8.23%[64], with the developed countries at the lower end. Most of the research on maternal obstetric morbidity has been retrospective, cross-sectional, hospital based, limited to specific population groups and using medical records; few have been prospective studies (Brace et al. 2004; Waterstone et al. 2001; Filippi et al. 1998; Geller et al. 2004 and Say et al. 2004).

In the UK according to HES Online Maternity Data, the most commonly reported complications of labour and deliveries for 2008-09 were perinatal lacerations (37.1% of all deliveries); fetal distress (21.8% of all deliveries); postpartum haemorrhage (10.1% of all deliveries) and long labour (10% of all deliveries).

There is evidence that PPH is one of the main contributors to maternal mortality and severe morbidity around the world (Callaghan et al. 2008; Zhang et al. 2005; Penney et al. 2007). PPH accounted for 11% of maternal deaths globally (Ronsmans et al. 2006) and there was a 1% fatality rate in 14 million cases of obstetric haemorrhage (Abou Zahr 2003). A prospective observational nationwide study from Scotland (Brace et al. 2004) and a case-controlled study from a defined population in South East England (Waterstone et al. 2001), found that at least two thirds (50% in Scotland) of all cases of severe maternal morbidity were related to severe haemorrhage and that obstetric haemorrhage remains a problem in terms of mortality and morbidity in UK (CEMACH 2007; Knight et al. 2007 UKOSS study). Ten of the 17 women who died in UK for the period 2003-2005 (CEMACH 2007) from haemorrhage, received less than optimal care, which raised questions about appropriate management and failure by staff to recognize certain conditions leading to haemorrhage. According to the CMACE (2011) – "*there has been a decline in mortality from PPH in UK over the last 50 years. PPH dropped to sixth place as the*

[64] Depending on the criteria in the studies to classify patients as being 'near miss' – disease-specific (common conditions, e.g. preeclampsia, haemorrhage); management-specific (related to response to disease, e.g. hysterectomy or admission to ICU); and organ-system dysfunction/failure based (dysfunction or failure related to each organ system).

cause of direct deaths in the last confidential enquiry from 3rd place in the 2003-2005 CEMACH". PPH is also included as a high risk condition in the Clinical Negligence Scheme for Trusts (2012)[65] and PPH (>1500ml) is in the suggested trigger list for incidence reporting in maternity (RCOG 2009). RCOG (2009/11) guideline defines two types of PP haemorrhage – primary (with blood loss of 500 ml or more, from the genital tract within 24 hours of the birth) and secondary *("abnormal or excessive bleeding from the birth canal between 24 hours and 12 weeks postnatally"*). PPH was reviewed as an outcome in *Towards Better Births* (HCC 2008). The review presented PPH as a maternal morbidity indicator and a risk marker for maternal mortality (the other two were eclampsia and transfer of mother to ICU). PPH had the highest incidence compared to eclampsia and ICU. Great variations in PPH across trusts were observed and there was no certainty if these variations were actual or data driven (there were inherent difficulties in measuring and reporting PPH, blood loss is poorly evaluated and often rely on subjective assessment). The data collected was for: *'significant'* blood loss (1000 ml or more), with incidence of 27 per 1000 births (for median trust with variations of 0.4-100 per 1000 births); and *'major'* blood loss (2500 ml or more), with lower incidence of 1.9 per 1000 births (for the median trust with variations of 0.1-8 per 1000 births).

Research by Roberts et al. (2008 and 2009); Brace et al. 2004 and Waterstone et al. 2001, estimated incidence and predictors of severe maternal obstetric morbidity and suggested ways of measuring it. The study by Roberts et al. (2008) created and validated a composite "*maternal morbidity outcome indicator (MMOI)"*, using NSW Australian routinely collected population health data (linked birth and hospital discharge data) which included both diagnoses and procedures (for being more reliably reported). The MMOI[66] was a broad measure of severe adverse outcomes (morbidities from the severe end of the spectrum and broader than '*near miss*'

[65] PPH is under Standard 3 (High Risk Conditions), Criterion 7.

[66] The list of the diagnoses (with ICD10 Diagnosis Codes) and procedures could be found in Appendix B of the article.

morbidities) associated with childbirth rather than causal conditions (i.e. haemorrhage, preeclampsia). It was developed to assess the quality of maternity care and as a result did not include factors that may be directly related to service provision (excluded certain procedures that may be directly related to service provision) or modes of care[67]. It identified women (both low and high risk pregnancies) affected by serious adverse events as a consequence of birth and frequent enough to inform potential suboptimal care. The researchers considered that the MMOI could be used to study factors predicting major morbidities; to evaluate effectiveness of obstetric interventions; to compare outcomes for women across obstetric levels of care and to estimate health system costs for major maternal morbidity (because costs are based on procedures).

The MMOI were used to examine the association between adverse outcomes and maternal, pregnancy and birth characteristics among women with postpartum haemorrhage (PPH), as well as the effect of adverse maternal outcomes on maternal length of stay (Roberts et al. 2009). They found that the adverse maternal outcomes increased by 21% for the period of 1999-2004, and that this increase was almost entirely among women who experienced a PPH (67% of the women with adverse maternal outcomes had an obstetric haemorrhage). Among women with PPH the increase in adverse maternal outcomes was 14%. The risk factors for those women included: maternal age, parity, previous CS, multiple pregnancy, smoking, medical conditions, antepartum haemorrhage, induction of labour and birth in small rural hospital. Augmentation of labour, perineal tears, episiotomy, regional analgesia and gestation were not predictive of adverse outcomes among women with PPH.

Other studies (Joseph et al. 2007; Ford et al. 2007) have found that the increase in PPH in some developed countries was not determined by changes in the risk factors mentioned above (maternal age, CS, multiple pregnancy etc.). These studies did not

[67] ICU, length of stay, readmissions, availability of beds, as well as 3rd /4th degree tears were not included in MMOI.

take into account certain obstetric practices, such as active management of third stage of labour and regular monitoring following delivery, which were known to be effective in reducing PPH (Prendiville et al. 2000; Lalonde et al. 2006). Inadequate monitoring after delivery for example was explained by staff shortages in small rural hospitals in Australian NSW by a survey investigating an implementation of PPH management policy (Cameron et al. 2007).

Observation is vital antenatally, during labour and postnatally to identify women at higher risk (those with multiple pregnancies, previous CS, renal or cardiac disease) of adverse outcome in case PPH occurs. Thus midwifery/obstetric staffing levels and skill mix may have an important role in prevention of PPH. In relation to all above, obstetric haemorrhage emerges as one of the possible outcomes to be investigated in relation to quantity and quality of obstetric/midwifery care.

In 2003 the AHRQ (Agency for Healthcare Research and Quality) at the Department of Health and Human Services in US created a list of patient safety indicators, including three related to obstetric trauma. AHRQ Patient Safety Indicators (PSIs) were "*measures that screen for adverse events that patients experience as a result of exposure to the health care system; these events are likely amenable to prevention by changes at the system or provider level"* (AHRQ 2003).

Three indicators of obstetric trauma were selected to represent cases of potentially preventable trauma during delivery:

- Cases of obstetric trauma (3^{rd} or 4^{th} degree lacerations) per 1000 instrument assisted vaginal deliveries;
- Cases of obstetric trauma (3^{rd} or 4^{th} degree lacerations) per 1000 vaginal deliveries without instrumental assistance;
- Cases of obstetric trauma (3^{rd} or 4^{th} degree lacerations) per 1000 Caesarean deliveries

After numerous empirical tests, the panel at AHRQ[68] concluded that all three indicators performed well regarding reliability, relatedness and persistence over time. However the reliability strength, measured by the so called '*signal ratio*' was different for the three indicators. '*Signal ratio*'[69] measured that part of the total variation across hospitals which was due to systematic differences in hospitals performance rather than random variation. The *'signal ratio'* for obstetric trauma indicator was:

- *moderately high*[70] (69.9%) *for vaginal deliveries with instrument,* indicating that the observed differences in the risk adjusted (age, comorbidities) rates "*likely reflect true differences"* across hospitals. The *'signal standard deviation'*[71] was also high, i.e. the differences between hospitals were explained mainly by hospital characteristics. And the *'signal share'*[72] was high, i.e. the hospital seems to be more important in accounting for the obstetric trauma rate while other potential factors were less important (like patient characteristics/co-morbidities).
- *high* at 86.4% *for vaginal deliveries without instrument,* suggesting *true differences* in risk adjusted rates between hospitals; the *'signal standard deviation'* was high, while the *'signal share'* was relatively low compared to

---

[68] http://www.premierinc.com/safety/topics/guidelines/downloads/Pt-Safety-psi-guide-v31-2007sm.pdf

[69] "Signal ratio (the proportion of the total hospitals variation that is truly related to systematic differences in hospital performance rather than random variation)".

[70] Always compared to other indicators.

[71] "Signal Standard Deviation (the systematic differences between hospitals are high/low and more/less likely associated with hospital characteristics)".

[72] "Signal Share (the measure of the share of the total variation - hospital and patient, accounted for by hospitals. The lower the share the less important the hospital in accounting for the rate and the more important other potential factors, e.g. patient characteristics)".

other indicators, suggesting that hospital characteristics may be less important than patient or other characteristics; and

- *low (compared to other indicators) at 45.9% for caesarean delivery,* suggesting that the differences in risk adjusted rates *may not reflect true differences* between hospitals. The *'signal standard deviation'* and *'signal share'* for this indicator were also lower than many indicators (possibly meaning that if the differences between hospitals were true differences, these were not so much due to hospital characteristics but more to patient or other factors).

Roberts et al. (2007) used the AHRQ methodology to investigate risk factors for maternal birth-related trauma and to compare national rates of $3^{rd}/4^{th}$ degree perineal lacerations with Iowa state rates. They suggested that a combination of risk factors related to conditions recorded for the woman and her infant and the procedure/instrument used all contributed to the obstetric trauma. The conclusion was that episiotomy could be a risk factor or protective factor for birth-related trauma depending on the risk-adjustment content and the type of trauma. Vacuum and forceps procedures were both related to all types of birth trauma and forceps particularly related to $3^{rd}/4^{th}$ degree lacerations. There was other evidence that birth trauma may result from the use of instrument, episiotomy, or may be a spontaneous event or a combination of factors (Renfrew et al. 1998; Beckman et al. 2005).

A systematic review by Hartmann et al. (2005) concluded that maternal outcomes of routine episiotomy were not better than those with restrictive use while Renfrew et al. (1998) suggested that routine episiotomy procedures were not justified by the empirical evidence, with a higher chance of intact perineum if avoided. There was also a reduced risk of maternal trauma if instrumental delivery was performed by vacuum than forceps procedures (Johanson and Menon 1999). Some of those risk factors/interventions (forceps, vacuum extraction, episiotomy) could be controlled or avoided. Others related to the womens' characteristics (age, race, co-morbidities, size of baby) could neither be controlled nor there was conclusive evidence of their relationship to obstetric trauma.

Physical birth trauma and particularly $3^{rd}/4^{th}$ degree perineal lacerations can be considered as an outcome potentially sensitive to quantity and quality of obstetric/midwifery care because of the indirect effect of interventions (episiotomy, forceps, vacuum extraction) on birth-related trauma. The decision to perform the intervention, the type of intervention and the timing of interventions may be determined by staffing levels, experience and the skills of the operator (Keriakos et al. 2013). There was evidence that some interventions increased the likelihood of co-interventions (the so called '*cascade of interventions*') to observe and prevent further adverse effects, i.e. the use of epidurals increases the likelihood of use of electronic fetal monitoring, intravenous drips, synthetic oxytocin, drugs for hypotension, vacuum extraction or forceps, episiotomy, prevents mobility during labour and predisposes to higher morbidity (Caton 2002). Investigating birth trauma and interventions in relation to staffing numbers and skills have to be controlled for by a well developed risk model (incorporating patient characteristics and co-morbidities).

3.8.3 OPTIMALITY

The concept of optimality was presented in the literature review as an alternative - it measures process and outcomes of maternity care, from the position of health rather than illness. An optimal situation was defined as*: "a birth without complications or interventions occurring at the proper time and resulting in a healthy baby and a healthy mother*" (Wiegers 1998; Prechtl 1980). Optimality measures optimal evidence-based events and not rare adverse events. The move away from using maternal mortality and morbidity events as outcomes was driven by the rarity of these events in high income countries and by the renewed initiative to measure positive events, i.e. 'normal' (after all pregnancy and labour for the majority of healthy women in the developed world are normal events which do not require interventions, unless there are clear indications for them).

Optimality is considered a narrower concept than normality. "*Normal" is frequently defined in health care as the absence of abnormalities or adverse events*" (Murphy

and Fullerton 2006). Optimality escapes the complexity in defining what is normal and what not - it means best possible conditions for mother and baby.

The concept of optimality was developed by Prechtl (1968, 1980) to identify infants with a "*perfect start in life*". The optimal condition applied to mothers in perfect health, who had no problems during pregnancy and birth; interventions at any stage indicated problems and that meant less than optimal situation.

The concept of optimality was adopted by Wiegers et al. (1996, 1998) for use in Netherlands to measure quality of midwifery care for low-risk women and subsequently to measure differences in obstetric outcomes for low-risk women with planned home and hospital births; by Murphy and Fullerton (2001, 2006) in US who developed and validated the Optimality Index (OI) US which was used to investigate differences in the process and outcomes of nurse-midwifery care for women at low/and moderate risk; by Cragin and Kennedy (2006) in US to compare midwifery and medical care practices for low- to moderate-risk women and to measure the optimal perinatal outcomes; and piloted by Sheridan and Sandall (2010) for use in UK as a new approach in evaluating outcomes of midwifery care (community-based caseload midwifery within a maternity unit in a inner city teaching hospital in England) for mixed risk women.

The Optimality Index (OI) incorporates two parts – a *Perinatal Background Index* which includes items related to the background risk profile of mothers (their social and medical background, obstetric past history, present pregnancy); and a *Perinatal Outcome Index* which contains optimal antenatal, intrapartum, neonatal and postpartum components. The items used to construct an OI should reflect the best available evidence (in UK mainly NICE and CEMACH recommendations) about obstetric care. According to Murphy and Fullerton (2006), OI-US for example was not a risk assessment or benchmarking tool; it did not measure risk or predict adverse outcomes for mothers, even though it used many risk factors and adverse events in its content. It aimed to measure the process and outcome of a normal pregnancy.

These studies provide evidence that the OI can be used by maternity staff to highlight areas in the process of care that were linked to undesired outcomes (Murphy and Fullerton 2006); there were no significant differences in optimal birth outcomes for the women under caseload and standard care, apart from a higher breast feeding rate among women in the caseload group (Sheridan and Sandall 2010); the women cared by certified nurse-midwife experienced more optimal care processes than the ones cared by physician, in terms of less use of interventions and technology; and that there were no differences in neonatal outcomes.

3.9 PROCESS/OUTCOME INDICATORS SELECTED FOR THIS THESIS

In this section, the indicators selected in this thesis for the modelling of the impact of staffing levels are described and explained.

Two process indicators were selected, caesarean section and instrumental delivery, and one outcome indicator, normal birth. They are each now described in detail.

3.9.1 CAESAREAN SECTION

3.9.1.1 OPTIMAL CAESAREAN SECTION

The increased rates of CS and the unjustified routine use on very healthy mothers in the developed world became a concern for policy makers, clinicians and childbearing women. WHO recommended rate for CS in 1985 was for 10-15% (WHO 1985). This upper threshold was a theoretical estimate at the time but several recent global studies (Althabe et al. 2006; Villar et al. 2006) have since supported that estimate. These studies found no decline in maternal and neonatal mortality and morbidity if the CS rate was more than 15% and even an increase in maternal and neonatal mortality and morbidity when higher than 15% (Villar et al. 2006). Real benefits were observed for rates up to 10%. There were no added maternal and neonatal benefits and even negative consequences for mother and baby if the CS rate was beyond a certain threshold (Betran et al. 2007; Althabe et al. 2006; Belizan et al.

2007; Villar et al. 2006; Barros et al. 2005). Rates higher than 15% may result in more harm than good (Althabe et al. 2006b) though further high quality research is still needed to confirm this.

In their 2009 handbook, WHO acknowledged the existence of a growing body of research showing the negative impact of a high CS rate; that both very high and very low rates were dangerous but that the optimum rate was unknown. It identified a lack of empirical evidence for an optimum percentage or range of percentages for CS (WHO 2009).

Lower CS rates are favoured by the medical profession as long as there is no risk of neonatal and maternal morbidities. It is also widely accepted in the developed world that lower CS are related to better clinical practice and performance (Fantini et al. 2006). The reasons identified in the literature for the increase of CS rates included changes in the socio-demographic factors (such as aging mothers, ethnicity, IVF procedures often resulting in twin pregnancies, obesity and maternal morbidity); women's preferences; technological advances; differences in obstetric attitudes and practices; training and experience of staff and maternity litigations, etc. A study by Loudon et al. (2010) using routinely collected data from a tertiary obstetric unit in London for a period of 10 years, considered factors such as relative safety of the CS procedure with the introduction of regional anesthesia, thrombo-prophylaxis and use of antibiotics. The authors also discussed that the combined reduction of hours training for junior doctors by two thirds[73] may have had an effect on experience and confidence of staff resulting in preference to perform a CS procedure instead of dealing with complicated labour (i.e. instead of instrumental delivery or after failed instrumental delivery).

[73] Junior doctors in training working hours have been reduced from more than 100 hours per week to 48 hours since 1993 and the introduction of Calman training programme in 1995, which reduced training from 12 years to 7 for surgical specialities including obstetrics).

In the United States, a report by Sakala and Corry (2008) revealed perverse financial incentives for caesarean birth (evidence for higher payments for CS than vaginal births in US) and the flexibility it provided for providers to plan for staffing levels and workload. A planned CS offered predictability and convenience - shorter procedure timing, advanced staff planning, weekday working hours for staff (scheduling births by time of day, day of week and non-holidays, which is also cheaper when outsourcing staff), quick turnover of delivery rooms and higher fees. There was also a suggestion for supplier-induced demand in US - a reconfiguration of the hospitals, resulting from the increased investments in operation theatres and in postnatal beds (for longer recovery) to accommodate the scheduled CS and induction procedures during the working weekday hours. This creates pressure to get returns to these investments. There is some evidence that the NHS Payment by Results (PbR) system introduced in 2003 also produced perverse incentives for NHS trusts in England to perform elective CS, as it created a greater profit margin against the tariff than normal deliveries did (NHS Institute for Innovation and Improvement, 2006). This was expected to be resolved by the proposed 2013/14 single payment system which aimed to promote normal birth activities and woman-centeredness and to stimulate proactive care and prevention rather than reaction (a Commissioner will pay a provider for the whole pregnancy-related care during pregnancy, birth and postnatally, but not for individual activities) (DoH 2012).

The main disadvantages discussed in the literature of overuse of the CS procedure included the higher costs, maternal complications and increased chances of repeat CS for consecutive births (9 out of 10 in the US, Sakala 2006), longer stays for recovery and more intensive care for women and their babies, surgical injuries and infections, emergency hysterectomy, blood clots and stroke, prolonged postpartum pain, PPH, difficulty in establishing breastfeeding and bonding with the baby and generally a long-term negative impact on maternal well-being (Sakala 2006). The higher likelihood of experiencing most of those adverse outcomes with CS compared to vaginal births were confirmed in a systematic review by Sakala (2006). The review also identified involuntary infertility and reduced fertility, poor mental health, low

birth weight, placenta previa, accreta and abruption, stillbirth, preterm birth and maternal death in future pregnancies and births for women who had a previous CS. There was also a risk of premature delivery as the estimates of fetal gestational age were frequently imprecise (Engle 2006). Clinically CS is recommended for prolapsed umbilical cord, placenta previa, placental abruption and breech baby position (NICE 2004).

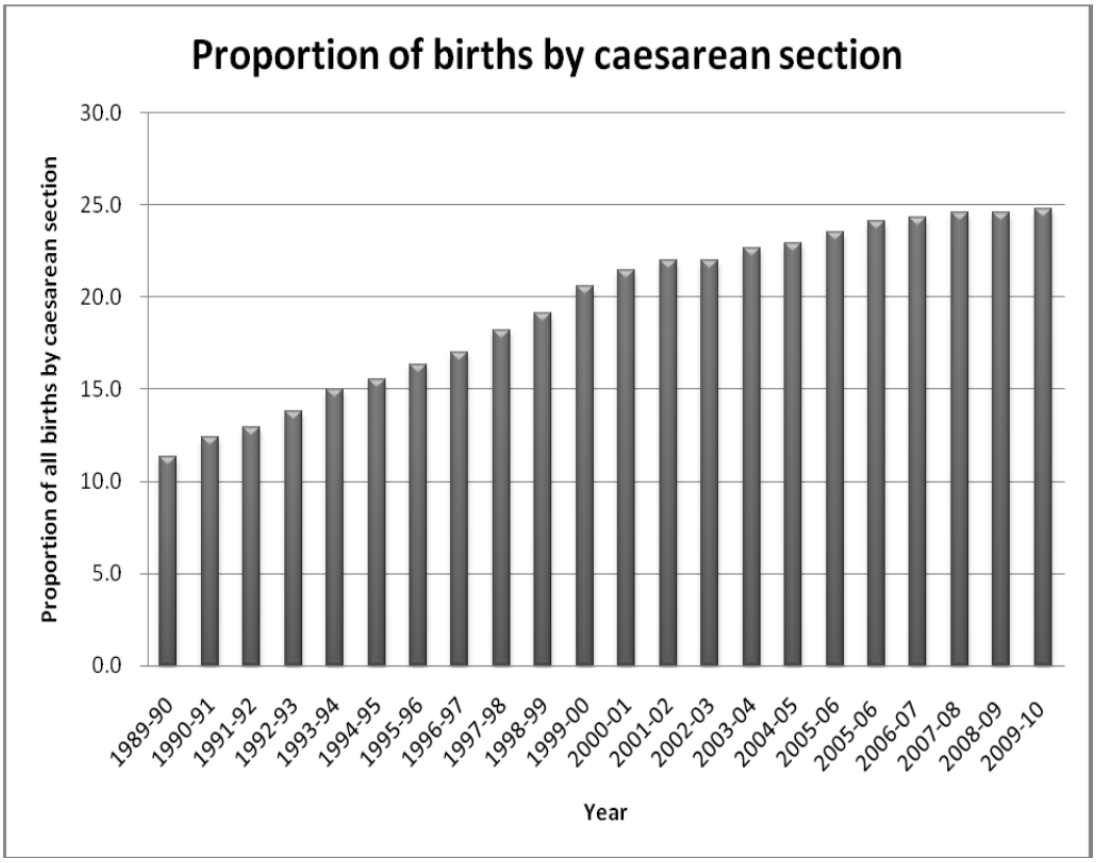
3.9.1.2 CS RATES IN ENGLAND

A National Sentinel Caesarean Section Audit in 2001 (England, Wales and Northern Ireland) was carried out in response to wide variation of CS rates across the UK. It showed that women with previous CS, preterm labour and breech baby were most likely to have given birth by CS. The main reasons reported for performing CS were: fetal distress (22%), failure to progress (20%), previous CS (14%), breech baby (11%) and maternal request without a medical reason (7%).

There were 668,195 deliveries in England in 2010/2011, of which 24.3% (24.8% in 2009/10) were CS deliveries - 9.8% elective and 14.8% emergency CS (HES online 2010/2011). In comparison the 1980 rate was 9%.

The review of the maternity services in England in 2007 (HCC 2008) reported that 60% of all CS were emergency CS, with trust variations of 40-70% and average length of stay after CS of 3.4 days vs. 1.4 days for normal vaginal births. Figure 1 below shows the annual share of births carried out by CS in NHS hospitals in England since 1989-90 (NHS Maternity Statistics 2009/10; NHS IC 2010). It shows a large increase in CS in the last 20 years with tailing off in the last 4-5 years.

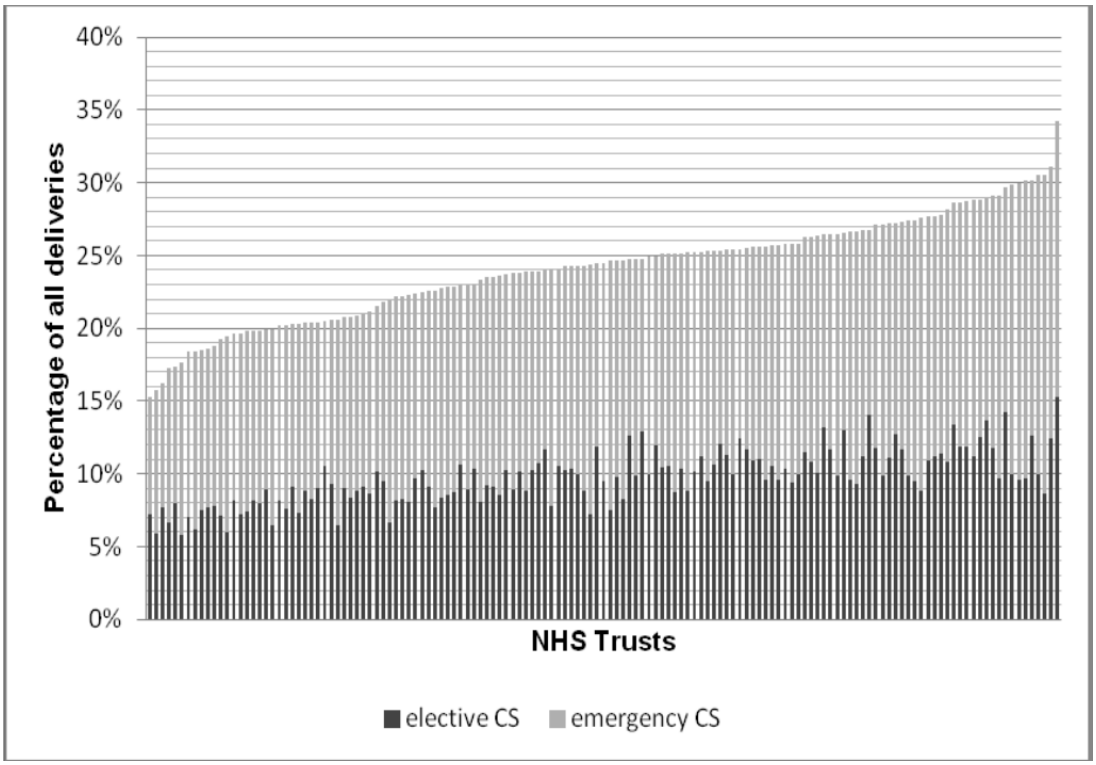
Figure 1: Proportion of births by caesarean section 1989-2010



Source: NHS Maternity Statistics 2009/10, NHS IC, 2010

Figure 2 below shows variations in elective and emergency CS across NHS trusts in England in 2009/10. It shows a variation in the overall CS rate of 15-34% (from 10-43% in 2000 in Paranjothy et al. 2005) and of planned CS from 6% to 15% (NHS Maternity statistics 2009/10; NHS IC 2010).

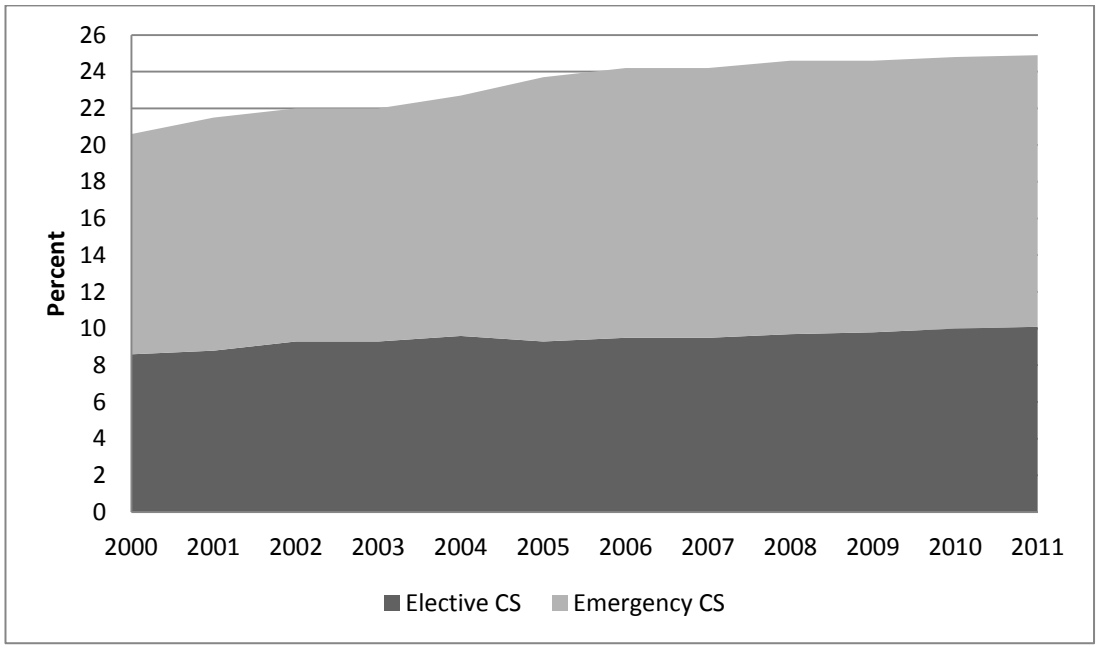
Figure 2: Variations in elective and emergency CS across NHS trusts in England



Source: NHS Maternity Statistics 2009/10, NHS Information Centre, 2010
http://www.hesonline.nhs.uk/Ease/servlet/ContentServer?siteID=1937&categoryID=1475

Figure 3 below shows the elective and emergency CS trends in England between 2000 and 2011 (BirthChoiceUK).

Figure 3: Trends in emergency and elective CS in England 2000/2011



Source: BirthChoiceUK

3.9.1.3 VARIATIONS IN CS RATES ACROSS TRUSTS IN ENGLAND

Results from a study by Joyce et al. (2002) suggested that the level of junior but not consultant medical staff was positively correlated with CS rates. There was no association between midwifery staffing levels and CS rates. However the study by Ashcroft et al. (2003) referred to earlier, observed several situations where shortages of midwifery staff resulted in risk situations for mother and baby as described previously.

Clearly clinical factors and patient characteristics play a major role in the women's method of delivery. However, risk-adjusted models of caesarean delivery have shown that the variations in CS rates persist after adjusting for these factors. A study by Bragg et al. (2010) presented evidence of wide variations in unadjusted (13.6% - 31.9%) and adjusted CS rates (14.9% - 32.1%) after adjusting for maternal characteristics and clinical factors across NHS trusts in England. The overall CS rate was 23.8% (9.3% elective and 14.5% emergency).

The persistent overall variation in CS rates was mainly due to variation in emergency CS but not elective CS rates. Nearly three-quarters (72%) of elective CS were performed for breech baby or a previous CS. Almost all women with placenta previa, abruption and breech presentation had a CS (compliant with NICE guidance, 2004). Differences in management practices and the lack of exact definition for fetal distress or shoulder dystocia (as well as difficulties in their diagnosis) were noted as possible explanations for the emergency CS variations. This was a cross-sectional study using routinely collected data (HES), examining singleton pregnancies and the effect of clinical factors and maternal characteristics on CS variations in 146 NHS trusts (all trusts with >1000 deliveries in 12 months) in England for 2008 (Jan-Dec). ICD-10 was used for diagnostic codes and OPCS, 4th revision for operative procedures from HES data and maternity tail in HES was used for deliveries specific information (it provided information for only 75% of all deliveries). Maternal characteristics included: age (15-44), singleton pregnancy, ethnicity, parity, deprivation. Clinical risk factors included: breech presentation, dystocia, fetal distress, previous CS, pre-existing and pregnancy related diabetes mellitus, pre-existing hypertension, eclampsia and pre-eclampsia, placenta praevia and placental abruption and preterm delivery. Multinomial logistic regression was used for three outcomes: vaginal delivery, elective CS and emergency CS. 51% of all women had no risk factors for CS, of them 4.9% (15 431) had a CS (29% emergency and 71% elective). Thus overall variations of adjusted CS between trusts were mainly due to variations in emergency CS. Slightly higher overall variations were observed in larger trusts (>4000 deliveries) and the risk adjustment dissolved the south-north divide in CS.

The conclusion was that differences in practice style among the providers may contribute to these variations.

Other studies have confirmed persistent variations after risk adjustment with the possible link to organizational factors, staff attitudes, women's preferences and cultural background (Habiba et al. 2006; Mossialos et al. 2005; RCOG 2013)

3.9.1.4 WHY THIS THESIS FOCUSED ON EMERGENCY CS

Given the lack of consensus on the optimum CS rate and evidence for wide variations in unadjusted and risk adjusted CS rates between trusts in England (HCC 2008; Bragg et al. 2010; RCOG 2013), it made sense to investigate CS and to explore maternity workforce and trusts characteristics contribution to what seemed persistent variations among maternity care providers. The interest in emergency CS was brought about by Bragg et al. (2010) study which suggested that the overall variations in adjusted CS between trusts were mainly due to variations in emergency CS. It was of interest to investigate whether differences in staffing levels could explain differences in emergency CS.

According to NICE (2011) guideline the classification of urgency in CS relate to immediate threat to the life of the woman or fetus; maternal or fetal compromise which is not immediately life-threatening; no maternal or fetal compromise but needs early delivery and delivery timed to suit woman or staff. It was recognised that the definition of emergency CS (resulting in differences in interpretation and coding) may have contributed to the observed variations in emergency CS rates across maternity units in England (RCOG 2013) as the definition of urgency covered a wide range of clinical situations.

A particular interest in this thesis became the variations in primary emergency CS rate for women aged 15-44 years, who were nulliparous, with a single, live, term birth (37 or more weeks of gestation). This relatively homogeneous group was appropriate to study the reduction of primary CS rates and should provide a better measure of a variation related to non-clinical factors.

3.9.2 INSTRUMENTAL DELIVERY

Operative vaginal delivery using forceps or vacuum extractor is an important obstetric practice. The goal of the operative vaginal delivery is to facilitate a spontaneous vaginal birth, by speeding up the delivery in the second stage of labour with minimum physical and psychological maternal and neonatal morbidities (Majoko and Gardener 2012; Keriakos et al. 2013). *"The second stage of labour is defined as the interval between achieving full cervical dilatation and the birth of the baby"* (Majoko and Gardener 2012). Operative procedures are used as an alternative to delivering by caesarean section if complications occur during labour. The advantage of delivering by forceps for example instead of CS was that women were more likely to achieve a subsequent spontaneous vaginal delivery compared to women with a previous caesarean section (Patel and Murphy 2004).

The majority of mothers and their babies seem to achieve safe and satisfactory outcomes after operative vaginal delivery (RCOG Green-top Guideline N26-2011a; Keriakos et al. 2013). However the forceps extraction for example could be potentially dangerous if performed by inexperienced obstetricians (Keriakos et al. 2013). Indeed the rise in obstetric litigations, though related to all areas of care often reflected immediate care on the labour ward, departure from clinical guidelines or inexperienced operators (Patel and Murphy 2004). The total value of claims between 2000 and 2010, related to operative vaginal delivery, were approximately £94 million (UK NHS Litigation Authority report 2012). Departure from safe practice is affected by the working environment (Vincent et al. 2002) and in case of operative vaginal deliveries could occur due to: failure to apply good medical judgement on the intervention needed; failure to evaluate risks correctly; failure to abandon instrument trial on time; failure to seek help from more experienced obstetrician; failure to adequately supervise junior obstetric staff members, etc. (Keriakos et al. 2013).

The recent RCOG (2011a) Green-top Guideline 26 has indicated operative vaginal delivery but only as guidance for presumed fetal compromise; maternal exhaustion; in nulliparous women when regional anaesthesia was used and there was a lack of

continuing progress for 3 hours (active and passive second stage labour) or for 2 hours when no regional anaesthesia was used; and to shorten second stage of labour for women with certain medical conditions such as cardiac disease (RCOG 2011a based on British Columbia 2001 Obstetric Guideline 14).

Certain factors during labour have been established in helping women to achieve vaginal delivery and reduce the need for operative vaginal birth. Most of the evidence is based on Cochrane reviews or meta-analyses and the techniques and factors were recommended in the RCOG 2011a Green-top Guideline 26. One factor was the continuous support of women during labour (Hodnett et al. 2013). Other techniques included: an upright position or tilting onto the left or right lateral position (Gupta et al. 2012); reduction in the use of epidural (Anim-Somuah et al. 2005); mobility during labour and in case of epidural use, delaying instrumental delivery after full dilation for 2 hours before active pushing for primigravida (Roberts et al. 2004).

Vacuum extraction is not recommended for gestational age less than 36 weeks; in cases of face presentation; and both forceps and vacuum extractor are contraindicated before full dilatation of the cervix (RCOG 2011a). Adequate assessment, choice of instrument, timing and whether to perform an operative vaginal delivery or not could be a complex decision which has to balance the risks and benefits of continuing pushing versus an operative delivery (RCOG 2011a) or the risks and benefits of instrumental delivery versus second stage caesarean section (Keriakos et al. 2013). The recommendations also include adequate preparation and planning which include full assessment; clear communication with the mother (verbal consent in delivery room or written consent for theatre deliveries); experienced and skilled operator; back-up plan in case of failure, i.e. theatre staff should be ready for a caesarean delivery in less than 30 min; anticipation of complications; availability of neonatal resuscitation professional (RCOG 2011a).

Trial of instrumental delivery is expected to be performed or supervised by experienced obstetricians to minimise the need for using sequential instrument or

risky second stage caesarean section. Essential recommendation by the RCOG (2011a) was that obstetricians should be first trained and experienced in spontaneous vertex deliveries before operative vaginal delivery training was initiated (RCOG 2011a).

Maternal obesity (BMI>30), macrosomia (>4000g), prolonged first and second stages of labour were among the few indications for possible failure of instrumental delivery (Murphy et al. 2001a). Failure of instrumental birth could lead to a difficult second stage caesarean section which carries an increased risk of severe postpartum haemorrhage; uterine and vaginal tears, hysterectomy and major maternal morbidity (Murphy et al. 2001a), while sequential use of instrument could lead to neonatal trauma (Murphy et al. 2003). Some of the unwanted maternal outcomes of instrumental delivery are perineal and vaginal tears (which can be reduced with experience but not completely eliminated). $3^{rd}/4^{th}$ degree tears were more likely with the use of forceps rather than vacuum (O'Mahony et al. 2010) and the reported rate was up to 7% after use of forceps (Keriakos et al. 2013). These can have long lasting consequences such as faecal and urinary incontinence, genital prolapsed and psychosexual problems (Keriakos et al. 2013). Training of operators was seen as critical for a successful instrumental delivery and for reducing rates of caesarean deliveries in the second stage of labour (Spencer et al. 2006). It is not known how many supervised procedures were necessary for achieving competence, or what was the best way to audit the performance once the operator was trained (rates of $3^{rd}/4^{th}$ degree tear was a potential measure) (RCOG 2011a). In addition in case of anticipated failure a trial of instrument was suggested to be performed in theatre (Majoko and Gardener 2012) by an experienced obstetrician to minimise the delays which could lead to maternal and fetal morbidity (Keriakos et al. 2013).

A study by Loudon et al. (2010) examined the trend in operative deliveries at full dilation for the period of 1992-2001 by a systematic anonymous audit of routinely collected data from a high risk obstetric unit in London. The study aimed to investigate factors associated with CS at full dilation in the context of evidence for increased risk of maternal and neonatal morbidity after CS compared to successful

instrumental delivery, both performed in second stage of labour, and the reduced training and clinical experience of junior doctors (traditionally intrapartum care was the preserve of the junior obstetric staff and midwives who focus on vaginal delivery when safe, while the pre-labour CS decisions were made by more senior obstetricians). The study established that the mean values of maternal age, gestational age, birth weight or factors such as fetal head position, failure to progress or fetal distress or use of regional analgesia did not change over a period of 10 years. However failed instrumental deliveries, mainly due to failed ventouse attempts, have increased over the decade (Loudon et al, 2010). This was possibly due to ventouse becoming the preferred instrument used by junior obstetric staff, replacing forceps. Further, direct CS rates without an attempt at instrumental delivery have increased as well as the CS in second stage of labour and at full dilatation due to both instrumental failure and reluctance to use instrument. There was a speculation but not certainty that the increase in failed ventouse; unwillingness to continue with second instrument (lack of confidence, experience and because of the morbidity associated with use of forceps); or complete reluctance to use instrument therefore opting for immediate CS were all consequences of less clinical skills, training, practice and availability of senior supervision.

A recent cohort study (Essex et al. 2013) of women giving birth in UK between 2000 and 2002, using UK Millennium Cohort Study (MCS) aimed to investigate maternal socio-demographic factors associated with unassisted vaginal birth, instrumental vaginal birth, emergency and planned caesarean section. 12 000 MCS maternal interviews were linked to hospital medical records with high level of agreement. The results were stratified by parity and only live births were considered. The association between a large number of self reported characteristics such as maternal age, ethnicity, educational level, social class, language spoken at home, migration status was explored with the mode of birth after adjusting for health, interpersonal, pregnancy, labour and infant covariates (such as fertility treatment, BMI, smoking, medical factors during pregnancy, interventions, complications during labour, birth weight and gestational age) in a multinomial logistic regression. The results showed

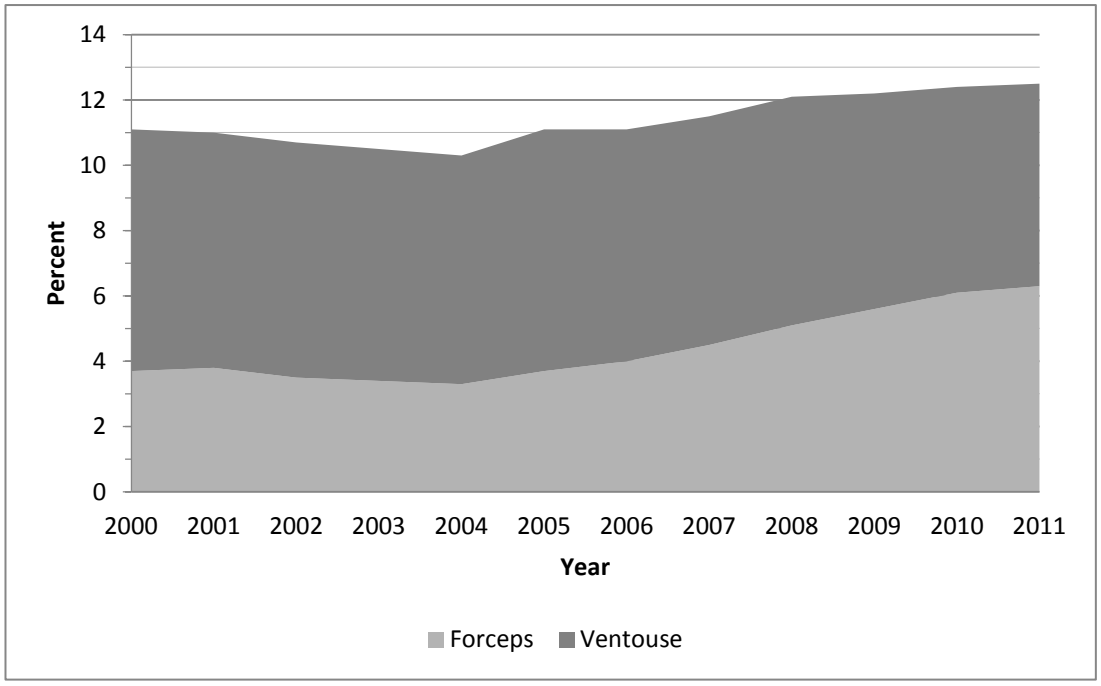
that the unplanned operative birth rate was much higher among first-time mothers (19.4% for instrumental vaginal births, and 19.1% for emergency caesarean section); the risk of emergency CS and instrumental deliveries rose with increasing maternal age; primiparous and multiparous Black mothers had the highest rate of emergency CS; primiparous Black women were less likely to have an instrumental vaginal birth while Pakistani or Bangladeshi women were less likely to have a planned or emergency caesarean section compared to White women; primiparous women in lower social class band were more likely to have an instrumental vaginal birth and an increased risk of a planned caesarean section compared to women in the highest social class band.

3.9.2.1 INSTRUMENTAL DELIVERY RATES IN ENGLAND

Operative vaginal deliveries in England have remained relatively stable between 10.3% and 13% over the last eleven years (Figure 4). Figure 4 was created using rates calculated by BirthChoiceUK, based on the routinely collected data from the English Hospital Episode Statistics (HES data) dataset. It seemed that in the first four years between 2000 and 2011 the use of forceps have declined (from 3.7% to 3.3%) but have gradually increased in the following seven years to 6.3% in 2011. Ventouse seemed to have been the preferred instrument over forceps in that period of 11 years (around 7%) but by 2011 the two instruments seemed to have an equal share in all operative vaginal deliveries (around 6.3% each).

Nevertheless a recent report from the National Audit Office (2013), regarding maternity services in England, established that the rate of adjusted instrumental births (forceps, ventouse or breech extraction) varied in 2011-12 from at most 9.5% in the lowest 10% of trusts to at least 16.1% in the top 10% of trusts, with nearly half of the trusts (69 out of 142) being significantly different (higher or lower) from the rest.

Figure 4 Instrumental delivery trends, England 2000/11



Source: BirthChoiceUk

3.9.3 NORMAL BIRTH

3.9.3.1 DEFINITIONS

World Health Organisation first addressed normal birth in a report published in 1996 (See Chapter 2.4). The policy and definitions adopted in England by the Maternity Care Working Party in their *Normal Birth Consensus Statement* and NICE (2007) guidelines recommendations were also presented in Chapter 2.4.

Dodwell and Newburn (NCT 2010) recommended normal birth as a measure of quality of maternity care and specifically midwifery care in terms of safety, effectiveness and positive experience for women. The authors made a note on the differences in normal birth terms and definitions; particularly the interchangeable use of normal birth and spontaneous vaginal birth.

"The term 'normal birth' is shorthand for a vaginal birth without any of the medical procedures that require hospital-based care, and are usually carried out by a specialist hospital doctor, including induction of labour, epidural or spinal anaesthetic, and the use of forceps, ventouse or caesarean section. The idea behind the term 'normal birth' is that it is the kind of care that can be provided either at home or in a birth centre by a midwife, though it is also possible in a hospital setting" (Dodwell and Newburn 2010:5).

Though spontaneous delivery was considered a useful audit indicator it was less relevant when measuring the overall intervention during labour (i.e. spontaneous vaginal birth is a wider term, it is the mode of the baby's emergence and could result after the use of induction of labour or epidural) (Dodwell and Newburn 2010). Normal birth, as adopted by MCWP, has a much stricter definition (see Chapter 2.4).

3.9.3.2 NORMAL BIRTH AS A MEASURE OF THE QUALITY OF MIDWIFERY CARE

Promotion of normality of childbirth is central to the quality of maternity services (RCOG 2008). Behind that recognition was the wide acceptance that childbirth is a normal life event (DoH 2004); and birth is a normal physiological process (Welsh Assembly 2002) for most women who are healthy and that with the right support majority of women were able to achieve straightforward vaginal birth with minimum interventions. Women also seemed to prefer to avoid interventions as long as the safety of their baby was not compromised and they were able to cope (Thomas et al. 2001; Green et al. 2003).

The following paragraphs are based on Dodwell and Newburn (2010) paper which provided evidence and discussion regarding adopting normal birth as a measure of the quality of midwifery care.

Several advantages for women who had a spontaneous vaginal delivery were listed (Dodwell and Newburn 2010). These included: reduced likelihood to suffer from pain after childbirth, compared to women who had caesarean section (NICE 2004) or assisted delivery (Johanson et al. 1993); increased likelihood to initiate breastfeeding

compared to women who had caesarean birth (NICE 2004); and a better start to motherhood (NCT 2002). In addition women reported higher satisfaction with spontaneous vaginal delivery compared to women who had interventions (induction and augmentation of labour; episiotomy or epidural) (Green et al. 2003).

There was evidence in the literature (Dodwell and Newburn 2010) to support various practices which enabled birth without interventions and promoted normal birth and positive experiences for women. These included amongst other: continuity of midwife-led care; offering birth at home or in a birth centre; one-to-one midwifery care in labour; provision of birth preparation classes; etc. The evidence behind these practices looked to establish the contribution of midwives in provision of safe and effective forms of care. Each of these practices was suggested (Dodwell and Newburn 2010) as individual measure of the quality of midwifery care. Few of them will be described below.

Continuity of midwife-led care

A Cochrane review (Hatem et al. 2008) found no adverse outcomes and no differences from midwife-led care in terms of fetal loss, neonatal death, low birth weight or admission to neonatal care. Midwife-led care increased the chance of spontaneous vaginal birth and reduced the chance of instrumental delivery; and resulted in reduced use of regional analgesia and episiotomies; women in addition were more likely to feel in control during labour and more likely to start breastfeeding. Women were also more likely to be cared for in labour by a midwife they have got to know, thus having a greater chance of continuity of care. Based on the results (including increased chance of spontaneous vaginal birth and reduced likelihood of instrumental birth) from the update of the above mentioned review (Sandall et al. 2013) the authors recommended that *"most women should be offered midwife-led continuity models of care"*(p.2) but caution should be applied in offering that advice to women with medical or obstetric complications.

One-to-one midwifery care in labour

The existing NICE (2007) clinical guideline recommends that women in established labour should receive supportive one-to-one care. The Cochrane review mentioned above (Hatem et al. 2008) did not identify any adverse outcomes from providing one-to-one midwifery care. Overall provision of continuous one-to-one support during labour reduced the need for interventions including caesarean, forceps or ventouse and epidurals (Beake et al.2001; Benjamin et al. 2001; Hodnett et al. 2013). In addition women who were supported during birth had a higher satisfaction and reduced feeling of trauma (Hodnett et al. 2013; Page et al. 2001; Ford et al. 2009).

Offering birth at home or in a birth centre

NICE (2007) intrapartum clinical guideline recommended offering women a choice of where to give birth, i.e. in a midwife-led unit, obstetric unit or at home (NICE 2007). In addition the NICE (2004) guideline on caesarean section suggested that planning a home birth reduces the likelihood of caesarean section for healthy pregnant women with uncomplicated pregnancies. Similar findings existed that planning for a home birth reduced the likelihood of instrumental delivery or an epidural (Chamberlain et al. 1997).

Two Cochrane reviews compared births in obstetric and ‘alongside’ or ‘co-located’ midwifery units (Hodnett et al. 2010) and home vs hospital settings (Olsen et al. 1998). The first (Hodnett et al. 2010) concluded that midwifery settings offered women a higher probability of spontaneous vaginal births, maternal satisfaction and less interventions. In addition no association between the setting and perinatal and maternal severe morbidity and mortality was observed, though the review lacked sufficient power to identify differences in rare adverse perinatal and maternal outcomes. The second review included only one randomised controlled trial with 11 women and found no differences in outcomes and safety between home and hospital settings.

A large scale, prospective cohort study (Birthplace 2011) of over 60,000 women in England, classified as healthy with low risk pregnancies at onset of labour, compared clinical outcomes for babies and women in different birth settings (planned births at home, in different types of midwifery units and in obstetric units). The study established that safety, low interventions and similar outcomes for babies were validated for 'low risk' women giving birth at midwifery units, with only half the rate of CS compared to the obstetric unit births. Women planning to give birth in all the three non-obstetric settings had lower odds of interventions, such as augmentation, ventouse or forceps delivery, intrapartum caesarean section, episiotomy. The adjusted[74] odds for 'normal birth' were also higher for the three non-obstetric settings. However the study warned about the higher risks for the babies of first time mothers giving births at home and for the higher transfer rates to obstetric units of first time mothers giving birth at home (nearly 50%) or in a freestanding midwifery unit (over a third). (See Place of birth studies in Chapter 3 for more details of the study).

3.9.3.3 NORMAL BIRTH RATES AND VARIATIONS ACROSS NHS HOSPITALS IN ENGLAND

According to BirthChoiceUK, 41.8% of women giving birth in NHS hospital in England in 2010-11 had a normal birth (Maternity Care Working Party definition). This rate was only 34.1% for first time mothers and 49.1% for women having a second or subsequent baby. The normal birth rate was about 42% for the period from 2004-05 to 2010-11. The unadjusted normal birth rate varied from 28.8% to 54% across NHS hospitals in England in 2010/11 (BirthChoiceUK website).

It will be of interest to investigate normal birth in relation to maternity staffing (particularly midwifery) because of its acceptance as a desired outcome for majority

[74] Adjusted for maternal age, ethnic group, understanding of English, marital or partner status, body mass index in pregnancy, index of multiple deprivation score, parity and gestational age at birth.

of healthy women; its policy relevance; as an outcome it is most related to the care provided by midwives; and because of existing variations across NHS trusts in England.

3.10 CONCLUSIONS

Quality of maternity care should be examined by indicators which include structure, process and outcomes as suggested by Donabedian (1980, 1988). Greener (1991) suggested that an instrument that includes aspects of each of these three indicators could be used to measure midwifery care.

Investigation of birth outcomes is a complex issue, as outcomes may be influenced by a multitude of factors. Kane's (1997) Model of Treatment and Outcomes for example ascertains that patient (age, ethnicity, deprivation, etc) and clinical characteristics (parity, co-morbidities, BMI, etc) have an impact on outcomes (mode of birth, complications, etc) directly and indirectly by influencing treatment decisions (i.e. oxytocin, electronic fetal monitoring, artificial rupture of membranes) (Hastings-Tolsma et al. 2009).

The literature has investigated a wide range of positive and negative outcomes with respect to interventions, models of care, staffing and place of birth. The emphasis was usually on negative adverse-outcomes.

For this thesis, two process indicators were selected, caesarean section and instrumental delivery, and one outcome indicator, normal birth.

These were preferred as a balanced choice: normal birth is considered as the only beneficial result regarding mother and baby. The others can be positive or negative, for example by saving life or implying some shortfall in the health system. Normal birth is more directly related to midwifery care, and the other two to medical obstetric care and therefore medical obstetric staffing skills and experiences.

Overall, the three indicators reflect the care provided by the four staff groups investigated in this thesis: midwives, doctors, consultant obstetricians and healthcare assistants. Further, variation in the three outcomes would be affected by women's characteristics, and the competence and experience of the staff providing care, and their skill mix within a particular organisation.

Thirdly, data on mode of birth are routinely collected at national level, though normal birth indicator was specifically derived for the analysis in this study. Data is therefore available, enabling an investigation of variations across trusts. Fourthly, they have each been the focus of previous research, allowing a comparison of the literature with this study.

And finally, they are policy relevant, because caesarean section is costly, and there has been a consistent rise in its rate over the past several decades. Normal birth is policy relevant as the preferred outcome for majority of healthy women, for which birth is considered a normal physiological process which should not be interfered with. Instrumental delivery is relevant regarding safety and adverse outcomes such as $3^{rd}/4^{th}$ degree tear and PPH.

4 CHAPTER 4 METHODS

4.1 OBJECTIVES OF THE CHAPTER

The aim of this thesis was to investigate the relationship between maternity staffing levels and selected mode of birth in NHS trusts in England for 2010/11. A cross sectional study was undertaken using routinely collected retrospective data.

The outcomes investigated were:

- Emergency Caesarean Section (CS);
- Instrumental Delivery ;
- Normal Birth (composite measure).

The Methods Chapter will explore in detail: the data sources for this thesis; quality issues including cleaning and manipulation of the data; and the modelling approach, including the selection of variables.

4.2 DATA SOURCES OVERVIEW

The main sources of data used were:

- Hospital Episode Statistics (HES) from 143 NHS trusts in England (including Maternity and Baby Tail) for April 2010/March 2011. All women delivering in an obstetric or maternity unit in an NHS trust in England;
- NHS IC Maternity Workforce Data – 2010-11 from the Health and Social Care Information Centre, for all maternity medical and non-medical staff in post, including bank staff;
- ONS Birth Registrations Data by Communal Establishment Code 2010/11;
- Department for Communities and Local Government, 2011 – for Index of Multiple Deprivation (IMD) 2010 overall ranking for each Lower layer Super Output Area (LSOAs) and Open Data Communities website for publicly

available UK postcodes with locations and links to LSOAs available from: (http://opendatacommunities.org/data/postcodes);

- The BirthChoiceUK[75] database provided information on NHS trusts and maternity units by location and type (trusts configuration data – Obstetric Unit/Alongside Midwifery Unit/Freestanding Midwifery Unit);
- Dr Foster[76] (for teaching trusts);
- London Local Supervising Authority[77], London LSA Annual report to NMC 2011 (for London trusts), available from http://www.nmc-uk.org/Documents/Midwifery-LSA-reports/LSAreports2010-2011/LSA_London_2010-2011.pdf;
- Monitor[78] - List of NHS Foundation trusts by authorisation date as at 1 April 2012 (for Foundation trusts status in September 2010), available from http://www.monitor.gov.uk/.

Access to most of these datasets and particularly HES was made possible due to a project funded by the National Institute for Health Research Health Services and Delivery Research programme, titled *"The efficient use of the maternity workforce and the implications for safety & quality in maternity care"* (Sandall et al. 2014, unpublished). The author of this thesis was a member of the study advisory board, attended meetings, and contributed to discussions, including implication of data

[75] BirthChoiceUK is a voluntary organisation helping women to make choices about their maternity care by providing information on their website www.BirthChoiceUK.com. It offers information and maternity statistics for each maternity unit in UK, which is accessible to parents and helpful in their decision on a place of birth.

[76] Dr Foster Intelligence is an independent healthcare information company and joint venture with the Information Centre of the NHS.

[77] London LSA is responsible for the statutory supervision of midwives in London in line with NMC's Rules and Standards and LSA guidance.

[78] Monitor is an independent regulator of NHS Foundation Trusts.

outcomes. Regarding the HES data, a project team member[79], executed consistency checks, cleaning and validation of the demographic and clinical data in HES core and maternity tail for 2010/11; developed algorithms and tested the NICE (2007) risk classification. HES inpatient records from 2000/01 to 2010/11; HES outpatient records from 2003/04 to 2010/11 and HES A&E records from 2007/08 to 2010/11 were stored in a MySQL database. The flat HES data files were reorganised into a relational database to speed up the data processing. Only the year 2010/11 was used at the modelling stage in this thesis. However, women giving birth were linked to previous inpatient and previous delivery records back to 2000/01, through HES_ID (the anonymised patient identifiers). This provided a better understanding of a woman's obstetric history, for example, it was possible to identify stillbirth or a caesarean in a previous pregnancy which were considered risk factors for the delivery recorded in 2010/11.

Other members[80] of the team constructed and verified the maternal risk composite measure (NICE 2007 risk)[81] described later; derived the normal birth composite outcome (explained further) and updated the NHS trusts configuration.

In addition the Office for National Statistics (ONS) provided the number of births by communal establishment code for 2010/11 under Open Government Licence v1.0. These were matched with communal establishment place using a list obtained from NHS Connecting for Health. Communal establishment place was matched to NHS

---

[79] Data cleaning methods, algorithms and the risk allocation method were initially developed by Rod Gibson Associates Ltd and were refined during the research project funded by the National Institute for Health Research Health Services Delivery Research programme (project number 10/1011/94) hosted by King's College, London.

[80] The mapping of the conditions listed in NICE (2007) Intrapartum care guideline to 4 digits ICD-10 codes was executed by Miranda Dodwell (BirthChoiceUK) and verified by Susan Bewley (Professor of Complex Obstetrics) and Rod Gibson (Rod Gibson Associates Ltd).

[81] NICE composite clinical risk measure was based on the National Institute for Clinical Excellence (NICE) 2007 guidelines and uses ICD-10 diagnostic codes and OPCS, 4th revision for operative procedures.

trust or (PCT) using the BirthChoiceUK database. ONS births were also used to identify the level of duplicate records in HES. The HES records were first restricted to NHS hospital deliveries resulting in a registrable birth[82] and duplicate delivery records were removed from the mother's records, while duplicates from the babies' birth record were kept. This process was executed by Rod Gibson Associates Ltd on behalf of the above mentioned project. ONS births instead of HES deliveries were used because ONS collects information on births and maternities (equivalent to deliveries in HES) and is considered the official and preferred source of that data (HSCIC 2011:8). Most of the information on live and still births is supplied by parents, who are legally required in England and Wales to register the birth of their baby within 42 days of the birth. The information is collected by registrars and supplied to ONS.

The Index of Multiple Deprivation (IMD) is produced by researchers at the University of Oxford for the Department of Communities and Local Government. The index was available for 2010 based on 2008 data and 2001 census information. The data are reported at Lower layer Super Output Area (LSOA), of which there are 32,482 areas. The LSOA with a rank of 1 is the most deprived, and 32,482 the least deprived, on this overall measure. The LSOAs could be mapped onto postcodes. All UK postcodes with locations and links to LSOAs were publicly available from Open Data Communities website (http://opendatacommunities.org/data/postcodes). The index measures relative deprivation in small areas across England. 38 indicators from 7 weighted domains are used to create the IMD 2010 overall ranking (Department for Communities and Local Government, 2011). The overall ranking does not provide much meaningful information and is usually converted into quintiles of deprivation which measure the relative ranking and allow for a non-linear relationship between deprivation and outcome variables.

[82] Not a registrable birth was defined if: the delivery was less than 21 weeks in the cleaned gestational age (GA); if GA was less than 24 weeks and a still birth was recorded in the cleaned birth status (BIRSTAT); and if an OPCS code for induced abortion (Q14) and an ICD codes (O00 to O08 for miscarriage and abortion) were found.

A special request (Data Sharing Agreement) to obtain the workforce data from the NHS IC was also submitted on behalf of the project, in which the author of this thesis was involved. The reason for this request was the need for more disaggregated data (headcount and FTE) for all medical and non-medical staff groups working in maternity services in England by: Agenda for Change pay band (or equivalent) and by O&G grade (medical staff), age, gender, ethnicity and nature of contract for each NHS Trust in England for the period 2006/7 – 2011/12. The data were encrypted and securely provided in Microsoft Office Access format.

The author of this thesis was not involved in these activities, but completed all of the data analyses presented in this thesis. The author of this thesis received a copy of the cleaned patient level 2010/11 HES data[83] (with ONS births by trust and IMD already added to it); HES data also contained the derived NICE (2007) composite risk measure and the composite normal birth indicator; a file with trusts configuration data and the original NHS IC workforce data files in Access format. The HES dataset was anonymised at individual level but contained unique identifier for each NHS trust. This thesis author's own work involved merging of all datasets at trust level; creating a hierarchical data structure; grouping and recoding of all variables used in the analytical part of this thesis; descriptive statistics analyses and subsequent multilevel modelling, all executed in IBM SPSS 22.

The following pages describe: data sources; data quality; the cleaning and preparation of the data, variables definitions and the multilevel logistic methodology.

4.3 THE ENGLISH HOSPITAL EPISODE STATISTICS (HES) DATASET

Hospital Episode Statistics (HES) is a retrospective routinely collected administrative dataset. It provides information on all admissions to the English NHS. It contains detailed patient level information. The collection of HES data goes back to 1989 for

[83] The provided cleaned patient level data did not contain the individual diagnoses and fields containing procedures women underwent.

inpatient episodes; 2003 for outpatient appointments and 2007 for A&E attendances (HSCIC 2011).

HES data are supplied by the NHS trusts. The data are sent to the Secondary Uses Service (SUS), as part of NHS Programme for IT, implemented by NHS Connecting for Health. SUS is a national data warehouse; the information is kept in a secure environment which guarantees patient confidentiality (HSCIC, http://www.hscic.gov.uk/sus). HES service subsequently provides cleaned and freely available aggregated data from SUS for all admissions to NHS hospitals in England. The data contains information on patients (age, gender, ethnicity, residence details and NHS number); administrative details (NHS trust identifier, dates and methods of admission and discharge, time waited, referral, GP, etc.), clinical information (diagnoses, interventions, procedures, operations, consultant specialty, etc.) and geographical information (where the patient was treated and the area where they live). HES records are cleaned and collated into annual datasets covering a financial year (1 April-31 March).

HES is a patient level dataset, but it records hospital activity not individuals (not numbers of patients), therefore some patients will have several records. Each HES record represents a Finished Consultant Episode (FCE), which is the total time a patient spends under the care of an individual consultant or another qualified health professional such as a midwife within one healthcare provider. The length of a FCE is derived from the epistart and epiend dates; admission and discharge dates give the spell duration. For majority of patients the stay in a hospital is presented by a single consultant episode. If the patient was transferred to another consultant or another hospital/trust, the patient will have more than one episode of care. This is more likely to happen if the initial admission was an emergency. These episodes could be aggregated into hospital spells (the total time a patient spends in one hospital); trust spells (the total time a patient spends in the hospitals of a trust); and a continuous in-patient spell – total time a patient spends in hospitals regardless of which trust. HES data are supplied without generated spells and there is no definitive methodology for constructing them (Sinha et al. 2012) but FCEs are "*sequentially numbered*" and

linkages across FCE and years for a patient are possible due to the *"patient's unique pseudo-anonymized identification number (the 'HES_ID') based on case note number, date of birth, sex, residential postcode and provider code"* (Sinha et al. 2012:5).

The clinical information in HES is recorded using *International Classification of Diseases* (ICD-10) diagnostic codes and *Office of Population Censuses and Surveys Classification of Interventions and Procedures* (OPCS-4) 4th revision. The ICD is used in health/death records to classify diseases and other health problems; for collecting national and international mortality and morbidity statistics; and as diagnostic tool in epidemiology, health management and for clinical purposes. "*ICD-10 was endorsed by the 43rd World Health Assembly in May 1990 and came into use in WHO Member States as from 1994*" (WHO, http://www.who.int/classifications/icd/en/).

OPCS-4 *is a "procedural classification for the coding of operations, procedures and interventions performed during in-patient stays, day case surgery and some out-patient attendances in the National Health Service (NHS)"* (http://en.wikipedia.org/wiki/OPCS-4). The OPCS classification was first released in 1987 as *Classification of Surgical Operations*. The 4th revision was implemented from 1992 named *OPCS Classification of Surgical Operations and Procedures*. NHS Information Authority (NHS IA) took over the responsibility for updating the classification in 1999 and NHS Connecting for Health in 2005. As from 1 April 2013 this responsibility was assigned to Health and Social Care Information Centre (HSCIC). OPCS is an alphanumeric nomenclature with a 4 character code system (first character is always a letter; second/third/fourth characters are numbers). Currently OPCS-4.6 revision is in use. But for the 2010/11 HES data OPCS-4.5 has been used. Each update reflects the advances in new procedures and becomes mandatory on 1st April of each revision year.

HES (though intended for financial reimbursement of providers' healthcare costs) has been extensively used to observe trends in NHS hospital activities and population

health trends over time; as a source of national indicators of clinical quality in assessment of effective care provision; to inform patient choice; for local planning and accountability and for Government policy evaluation (HSCIC 2012).

4.3.1 MATERNITY TAIL

HES maternity statistics have been released annually since 2000/2001. Information on admitted pregnant women is collected in their inpatient record, which changes to maternity record once they have given birth, and once updated is submitted by the local NHS providers' patient administration systems (PAS), via the Secondary Uses Service (SUS) (HSCIC 2012:8).

HES maternity information is collected in the maternity tail, which contains delivery (mother) records presented as delivery episodes and birth (baby) records as well as specific information on deliveries. A limitation of the maternity tail is that mother and baby records are not linked. Delivery episodes relate to number of mothers, not number of births (multiple births are counted once), but mothers delivering twice in a year are counted twice. Home births and private hospital births are not recorded in 'maternity tail' of HES. When the data are of good quality (accurate and complete), the HES maternity tail has the advantage of providing detailed antenatal, intrapartum and postnatal information, such as the gestation week of first antenatal visit, method of delivery, method of onset of labour and anaesthetics used, lengths of antenatal and postnatal stay, clinical details on mother and baby (diagnosis, treatment), birth status and birth weight; and organisation information (place of delivery and person conducting delivery).

4.3.2 QUALITY OF HES AND MATERNITY TAIL DATA

NHS providers submit their data to SUS and ultimately the accuracy of the released HES depends on the quality of the data submitted. NHS providers are required to submit accurate data as this is linked to being correctly paid for their activities. The data submitted to SUS is audited by the Audit Commission to ensure that NHS

providers receive correct payment from Payment by Results for the care they provide (HSCIC 2012).

Collection of HES data was intended for financial reimbursement of providers' healthcare costs. However, over the years, it has been extensively used in studies of the quality of healthcare provision, probably because HES as any routinely collected data, was readily available, relatively inexpensive to obtain, and covered large populations (Iezzoni 1997a; Ayanian 1999). The risks of using routinely collected data for healthcare quality assessment and comparison across providers relate to the differences in the quality of the data itself (by providers and year to year); the discrepancies in definitions and clinical coding between providers and its retrospective nature (Iezzoni 1997; Sinha et al. 2012).

The quality and completeness of HES data has been improving gradually over the last 10 years. Some studies suggested it was suitable for healthcare quality assessment (Aylin et al. 2007; Aylin-Bottle et al. 2007; Garout et al. 2008; Holt et al. 2012), while others expressed concern about its coding accuracy and completeness (Williams et al. 2002; Scott et al. 2011). Johal et al. (2013, see further below for details) suggested that variations existed in the accuracy and completeness of HES between clinical specialties; quality varied in relation to data for acute or long-term conditions; and for data items needed to measure process or outcomes of care.

A systematic review by Sinha et al. (2012) aimed to evaluate the methodological quality of published English medical studies assessing healthcare outcomes using HES dataset. The purpose was to produce guidelines for researchers using HES data in the future and to evaluate whether the quality of the studies had improved over time. It considered 148 studies (from 1989 to 2011), of which majority focused on surgical specialties and main themes related to inequalities and variations in treatment or outcomes. Multilevel modelling was used in 18.7% of the studies. The issues, critically examined in the review, related to data coding accuracy and completeness; methodological quality; study periods; inclusion and exclusion criteria; the use of ICD and OPCS codes; case validity; studies using mortality as an

outcome, (i.e. HES data linkage to ONS mortality statistics was investigated); the importance of defining spells for outcomes such as length of stay; how missing/invalid data and duplicate records were handled; the choice of covariates and the justification for risk-adjustment and interaction effects. The review concluded that over time, studies using HES data and related to healthcare outcomes have increased in volume, scope and methodological complexity.

The increased use of HES was possibly due to the fact that prospective studies or randomised control trials could be difficult to implement when assessing certain healthcare outcomes. Two of the reviewed studies (systematic reviews by Campbell et al. 2001 and Burns et al. 2012) showed high levels (>80%) of accuracy of ICD and OPCS diagnostic coding in HES over time, while another study (Scott et al. 2011) warned that variations in healthcare outcomes across providers were confounded by between provider variations in coding quality. Based on the deficiencies identified in the analysed studies, Sinha et al. (2012) provided a list of reporting and publishing guidelines for researchers using HES. The list is presented in Appendix IV and the recommendations were followed when cleaning HES data.

A recent report from the Royal College of Surgeons in England, RCS (Johal et al. 2013) assessed the use of HES for revalidation of procedures and their outcomes related to three areas: ischaemic heart disease, urological malignancies, and peripheral vascular disease, with the belief that findings could be extended to other surgical and medical specialties. The project also aimed to examine procedure-specific and disease-specific indicators that could be used to measure performance across hospitals and individual physicians. The project conducted a systematic review and 35 English studies (since 2002) were selected. All the studies used HES as their main source of data and investigated procedures and outcomes in the three specified areas. The majority of the studies presented procedure-specific results (none consultant-specific and rarely disease-specific); the prevalent topic was the relationship between hospital volume and outcomes; the most common outcomes examined were length of stay and in-hospital mortality; majority of the studies used age and gender for risk adjustment; most studies used simple descriptive statistics

and few used funnel plots to present hospitals specific results (mainly using annual hospital volume of procedures).

Case studies using HES data were also presented by the authors to revalidate indicators defined by the surgical specialist societies. Indicators were assessed for *"validity", "statistical power", "fairness" and "adequacy of coding specification"* (p.4) and just few were found "*fit for purpose*". National clinical databases from 2010 (related to the three areas) were also compared to HES and only two were found to have high enough "*completeness*" and "*case ascertainment*" to be judged against HES. The resultant comparison between HES and the two clinical databases showed a good consistency in procedural and diagnostic coding in HES for some indicators (abdominal aortic aneurysm repair surgery); the suitability of HES to measure outcomes (including mortality) from CABG (coronary artery bypass graft) procedure; and a recommendation for not using HES-derived indicators for percutaneous coronary interventions (PCIs) was made. The authors concluded that the decisions for using HES data when developing new indicators to evaluate performance should be individually made for each specific indicator, clinical area and condition/procedure involved.

The study by Johal et al. (2013) made few recommendations. One of them was for assessment of individual trusts coding practices when comparing performance across NHS trusts. In assessment of individual performance, indicators should be linked to the individual consultant, though this is problematic when care is provided by multidisciplinary teams and responsibilities are shared; case-mix adjustment was vital for disease-specific indicators (which measure the effect of all clinicians and teams involved on the outcome of treatment along the disease pathway). HES must be linked to external datasets which record patients at the time of diagnosis and over time (for the three areas reviewed it was difficult to judge the nature and the severity of patients' conditions from HES data alone as HES had the advantage of recording well the treatment procedure but not the timing of the diagnose) and correct level of analyses should be chosen, particularly the indicator which measures aspects of care

or outcomes should be linked to a unit that has most control over that care or outcomes (Johal et al. 2013).

This study was mentioned here in detail because some of the issues raised relate to the limitations of using matched patient level HES to trust level workforce data in the analyses of birth outcomes (see Chapter 7 for discussion of study limitations).

A study from Murray et al. (2012) examined methodological and processing issues in matching birth records to other healthcare records for the purpose of building population-level birth cohorts. The authors investigated the completeness of HES birth records (proportions of missing data in all *"baby tail"* fields) and found that all birth fields had improved between 2005/6 and 2009/10, but remained highly variable between NHS hospitals in England. The authors also presented a comparison (from 2007/08 HES) between 71 hospitals with high completeness of records (>=90%) and 85 low-coding hospitals on key *"baby tail"* fields (gestational age, birth weight) by hospital characteristics (mean number of births, maternity beds, specialist neonatal care facilities) and by maternal characteristics (mean maternal age, proportion of babies of non-white ethnicity and the proportion of babies in the most deprived Carstairs quintile). They concluded that the two groups of hospitals were similar (on mean values, though the data completeness on these characteristics was highly variable in both groups) and that using results from hospitals with high completeness of data may be generalizable and representative of all hospitals, thus recommending that when key birth information such as on gestational age and birth weight is missing, it may be preferable to select only hospitals with high levels of data completeness in these fields.

A comparison between all births in HES and ONS births registration over the 5 years period was also performed and it was reported that HES contained 87% of all live birth recorded by ONS. Among the reasons for this discrepancy was that births outside NHS hospitals may not be recorded in HES. The authors concluded that it was possible to create longitudinal cohorts which link individual birth records in HES to subsequent hospital admissions and other health records (Murray et al. 2012).

A study investigating linkages with HES maternity data showed high rates of linkage between HES and ONS births (Abrahms et al. 2002), while others linking the NHS Numbers for Babies (NN4B)[84] and HES birth records suggested that improvements of the quality of HES maternity data was essential (Dattani et al. 2011; Hilder et al. 2007). Birth data collected in the Millennium Cohort Study[85] was also validated against HES maternity data (Dezateux et al. 2006).

A recent study by Knight et al. (2013) examined the completeness and internal consistency of method of delivery recorded both in HES *'maternity tail'* and the procedure codes (OPCS) in the core fields of HES (this was done at national level and by NHS trust in 2009/10 HES). The purpose was to investigate how discrepancies in HES may misrepresent the information on maternity services, which derive their maternity statistics from it.

Given that a range of quality indicators for comparison of maternity services across providers use method of delivery from HES, the authors were concerned about the lack of information on the quality of this field. As the method of delivery was recorded in two ways: in the *'maternity tail'* (in *'Delmeth'* field by extraction from the electronic maternity information systems, where information is entered by midwives) and in the procedure codes (OPCS R17-25 from the core HES supplied by clinical coders from discharge notes), it was not clear which information was more complete and accurate. The overall level of coding agreement at national level was checked with kappa (k) statistics, while funnel plots were presented to examine variations in the coding consistency across trusts. Five different analyses rules for managing inconsistent data in delivery method were examined and their impact on three maternity statistics (emergency CS rate; rate of third and fourth degree perineal

[84] NN4B Service was introduced in 2002, it allocates a unique NHS number to each baby and collects limited information but some that is missing in birth registrations, such as gestational age.

[85] Millennium Cohort Study is a UK longitudinal observational cohort study of around 19,000 babies.

tears after instrumental delivery and elective caesarean section rate for breech presentation) which use method of delivery either in the numerator, denominator or both was tested.

In terms of completeness, the results (Knight et al. 2013) showed that for all 629,049 singleton deliveries in 151 English NHS trusts, 86.7% of the records had method of delivery recorded in both the 'maternity tail' and procedure fields; and overall method of delivery was much better captured in the procedure codes (available for 99.2% of records; all but 4 trusts had a procedure code entered for 95% of their deliveries) than by the 'maternity tail' (method of delivery was available in 87.5% of the records and only 96 out of 151 trusts had a code entered in more than 95% of their deliveries).

In terms of coding consistency, there was overall high degree of agreement between method of delivery codes in both 'maternity tail' and procedure codes when available and a divergent coding practices in minority of trusts (11 out of 136).

Knight et al. (2013) concluded that at national level the derived maternity statistics will not be affected much by the different rules of handling inconsistent data but maternity statistics derived at organisational/trust level was more sensitive to the different rules due to different levels of data completeness and coding in trusts. The study focused only on internal consistency checks of method of delivery in HES and did not validate a sample of HES against hospital medical records or a national birth register.

Inconsistencies in coding were largely due to emergency and elective caesarean section[86] though they both had high levels of consistency between maternity tail and procedure codes, which supported Bragg et al. (2010) conclusion that coding errors

[86] Inconsistencies in coding between emergency and elective CS contributed most (39%) to the overall coding disagreement in method of delivery, followed by inconsistencies between instrumental and non-instrumental vaginal births (19%) and inconsistencies for the type of instrument used (9%).

were unlikely to explain the large variations in emergency CS rates across NHS trusts. To an extent the inconsistencies in coding were possibly due to the definitions of emergency and elective CS used in the OPCS codes and *'maternity tail'*.

The authors (Knight et al. 2013) also concluded that method of delivery in HES can be used in national maternity statistics because of its high level of recording consistency. At trust level, majority of NHS trusts had high consistency levels and this provided evidence for the use of HES derived quality indicators for comparing performance across trusts. However trusts with diverging coding practices had to aim at improving the quality of their data due to risk of being incorrectly considered as outliers on performance indicators because of errors in the data provided by them.

Table 4.3-1 shows improvement of the data covered by *'maternity tail'* in HES for the years 2009/10 and 2010/11. For the main categories relevant to this research it shows that in 2010/11 method of delivery had 98% valid records, while there were 16% missing records on gestation length, 12% missing records on birth status, 11 % missing data on birth weight, 12% missing data on method of onset of labour and 15% missing data on anaesthetics used before or during delivery.

Table 4.3-1 Number of valid records in HES by maternity key fields, 2010-11 and 2009-10

HES maternity key fields	2010-11		2009-10	
	Number of valid known deliveries/records	% of valid known deliveries/records	Number of valid known deliveries/records	% of valid known deliveries/records
Place of delivery	577,670	86%	565,182	87%
Person conducting delivery	579,689	87%	562,702	86%
Reason for change of delivery	538,480	81%	525,138	80%
Intended place of delivery	577,514	86%	559,849	86%
Anaesthetics used before or during delivery	568,955	85%	547,912	84%
Method of onset of labour	586,744	88%	579,545	89%
Method of delivery	654,059	98%	635,741	97%
Duration of antenatal stay	591,744	89%	578,832	89%
Duration of postnatal stay	591,209	88%	578,340	89%
Gestation length	561,075	84%	522,088	83%
Gestation period in weeks at first antenatal assessment date	479,640	72%	464,557	71%
Birth status	585,527	88%	571,747	88%
Birth weight	592,845	89%	580,050	89%
Number of babies (parity)	608,874	91%	597,735	92%
Total deliveries	668,195	n/a	652,377	n/a

Source: Table 1, p.10 from https://catalogue.ic.nhs.uk/publications/hospital/maternity/nhs-mater-eng-2010-2011/nhs-mate-eng-2010-2011-rep.pdf

A list of core issues related to the quality and coverage of HES maternity data were presented in Murray et al. (2012) as well as in the annual bulletins by the NHS IC on HES maternity data (HSCIC 2011). They do overlap to an extent. A list of these issues is presented in Appendix IV, Box 2.

4.4 CLEANING OF HES

The cleaning of HES data followed the recommendations by Sinha et al. (2012), see Appendix IV.

4.4.1 MULTIPLE RECORDS

When multiple delivery records within some deliveries were not identical, one record had to be chosen to represent the delivery. This was achieved by first arranging all mother's registrable[87] delivery records into chronological order. When two successive records were separated by more than 30 weeks, they were considered to relate to different delivery episodes (early miscarriages would have already been removed, see definition of registrable birth in footnote). When two successive records were separated by less than 30 weeks and both records contained either baby sex, baby weight or gestational age and these were different, the two records were both kept as they were considered to relate to different delivery episodes. When multiple records related to the same delivery were identified, the best one was retained by using a scoring system based on most useful data with greatest relevance to the project (see Notes on p.155-56). The rest of the records which were not chosen were excluded. This cleaning procedure did not affect multiple births, as these generate a single delivery record.

4.4.2 GESTATIONAL AGE

Gestational age (GA) was used to determine registrable births (see footnote on previous page). GA inconsistencies related to some trusts providing GA in days rather than weeks, of which HES system truncated the last digit (of days) and the

[87] Not a registrable birth was defined if: the delivery was less than 21 weeks in the cleaned gestational age (GA); if GA was less than 24 weeks and a still birth was recorded in the cleaned birth status (BIRSTAT); and if an OPCS code for induced abortion (Q14) and an ICD codes (O00 to O08 for miscarriage and abortion) were found.

resultant HES data contained records for full term births with GA of around 27 weeks. Implausible GA values were set to unknown: 1) for all trusts when GA was between 10 and 20 weeks and the birth weight exceeded 1kg; 2) for trusts with more than 200 deliveries per year with recorded GA, if more than 10% of these records had GA ≤ 30 weeks then all records in that trust were set to unknown if the birth weight was not known to be ≤1kg.

4.4.3 PARITY

Parity is recorded in NUMPREG in HES. Parity was either not reported or reported incorrectly in HES. Some trusts seemed to have had too many first time mothers while others too many second time mothers. It was decided that trusts which had percentages of nulliparous women outside the range 20%-70% had all their parity data set to 'unknown' for that year. Parity was cleaned in stages by first using the criteria above to assess if a trust had submitted correct parity, and then woman's maternity history and ICD codes were searched to estimate unknown parity or correct existing parity.

4.4.4 MULTIPLE BIRTHS

Multiple birth deliveries were determined by using both NUMBABY[88] and ICD codes[89] with a denominator made of all usable records for which multiple births was known.[90] It was more likely that a multiple birth will be miscoded as singleton rather than the other way round, so a singleton was defined as not a multiple, when the plurality was known from either NUMBABY or the ICD codes.

---

[88] NUMBABY – if number of babies in a given delivery was greater than 1, i.e. (2,3,4,5,7,8); 6 was excluded as there was an excessive numbers of sextuplets in previous years.

[89] ICD codes from Z37.2 to Z37.8 (Z37.0 single live births; Z37.1 single still births; and Z37.9 unspecified outcome of delivery were excluded).

[90] Multiple births were known if NUMBABY was known; and ICD Z37 (excl. Z37.9) and Z38 (singleton/twins/other multiple by place of birth) codes exist.

4.4.5 MATERNAL AGE

Maternal age was derived from date of birth and episode start date and presented as patient age in whole years.

In the 2010/11 HES data of 657,480 births (used in the models, see Chapter V), there were 0.3% missing or set to unknown records for maternal age mainly because of miscoded date of birth. These included women reported as having given birth and aged: 60-120 years (small number), >400 years (due to a default age set to 16th century and used to calculate woman's age in some trusts), 0 years (due to recording current year as year of birth, few hundred) and aged 1-11 years (assumed to be miscoded and only few). Women aged 50-59 years (some assumed to be miscoded) and >12 years (assumed to be real) were left unchanged.

4.4.6 WOMAN'S ETHNICITY

HSCIC (2011) warned that ethnicity in HES should be used with care, as it may not be robust enough to support routine analysis of ethnic differences, although it was useful for addressing issues of data quality. Ethnic category from 1st April 2002 uses the definitions from 2001 census, thus differs from the ethnic group category collected between 1995 and 31st March 2002. Ethnicity was considered a 'soft' data item as patients self-select their category from preferred list and some patients choose not to provide this information (HSCIC 2011). There are issues related to the incompleteness of data collected centrally and to quality of ethnic coding HSCIC (2011). Given that in 2000/01 HES used different system of recording ethnicity, some of these old codes can still be found in the later years of HES. These "older" codes were mapped onto the newer ethnicity categories. If a woman's ethnicity was 'unknown' or 'not stated' both her inpatient and outpatient history was searched for valid ethnicity. This was successful 60% of the time. No attempt was made to make ethnicity consistent when it varied across a woman's history. Thus in the 2010/11 HES data with 657,480 deliveries (used in the models, see Chapter V), there remained 4.5% missing or unknown records of woman's ethnicity.

4.4.7 BIRTH WEIGHT

Baby's weight is recorded in grams in HES. Babies weighing in excess of 7kg; weighing less than 450g for recorded live births and less than 250g for recorded stillbirths were set to unknown.

4.5 NHS IC WORKFORCE STATISTICS, ENGLAND: 2010-11

NHS workforce statistics in England are collated by the NHS Information Centre (IC). Using mainly their administrative systems (Electronic Staff Records - ESR; NHS Payroll systems), the workforce data are submitted by around 400 NHS organisations, and on behalf of around 8,200 GP practices. The NHS Workforce Census is published annually, based on three censuses, and records the NHS workforce in England as at 30th September each year. The Census results are released in March of the following year. NHS IC also publishes timelier monthly workforce statistics. The responsibility for the accuracy of the data lies with the providers, while the IC uses Workforce Validation Engine (WoVEn) to check and validate the data against previous years' census figures for each organisation; consequently they report errors back to trusts and assign quality scores. Despite efforts being made to present workforce data accurately, there seem to be a degree of uncertainty about the actual size of the NHS workforce. The reasons mentioned by the IC are: the sheer size of the NHS workforce (1,431,557 NHS staff in September 2010), the constant reconfiguration of the NHS services and workforce composition, and the timing and ways the data is recorded at local levels (NHS IC 2011). NHS workforce data are also published by the Office for National Statistics (ONS) for the whole of UK as part of their Public Sector Employment Survey. The NHS IC contributes to that data collection on a quarterly basis.

Before 2010 trusts used annual census workforce surveys and supplied the data to NHS IC. From 2010 the workforce data is extracted directly from the Electronic Staff Records system into a central Data Warehouse, where it is updated on a quarterly basis.

A new methodology for the NHS staff headcount[91] was introduced in September 2010 Census. As a result comparison with previous years is not straightforward. The rationale for the new methodology was to remove duplications in the count for staff who work in more than one role[92] or contract. With the new methodology, non-medical staff who work in two or more part-time jobs in more than one organization or area of work are counted only once in each area of work or organization in the 2010 Census. Instead of counting all contracts and job roles in different trusts it now counts individuals. Thus a better match is achieved between the non-medical staff numbers from the Census and the national figures which count each staff member once. Overall, for non-medical staff the annual census has more precise count of absolute staff numbers at a National level, while the headcount figures for medical staff are more precise count of absolute staff numbers at an Organisational level (NHS IC March 2012a).

"Staff" relates to professionals "*holding permanent paid and/or honorary appointments that involve a degree of clinical work in the NHS hospital services and community health services"* (NHS IC March 2012a).

When statistics is presented as full time equivalents (FTE), *"numbers of part-time staff are converted into an equivalent number of "full-time" staff by taking account of the weekly number of hours or sessions in their contract... The maximum FTE for all staff is 1.26. All staff working less than full-time have a FTE less than 1"* (NHS IC March 2012a).

Table 4.5-1 below presents statistics on head count (HC) and full time equivalents (FTE) for selected staff groups (consultants in obstetrics & gynaecology (O&G);

[91] "Headcount refers to the total number of staff in either part-time or full-time employment within an organisation and/or area of work… FTE is the full time equivalent and is based on the proportion of time staff work in a role" (NHS IC March 2012a).

[92] "Role count is the total count of specific roles within an organisation and some people may have multiple roles either within or across organizations" (NHS IC March 2012a).

registrars in obstetrics & gynaecology (O&G) and midwives) in maternity services from the NHS IC September 2010 Workforce Census. The 2009 figures have been adjusted to allow for one-to-one comparison with 2010, while the 2000 figures have not been adjusted but according to the IC bulletin the effect of the new methodology on the actual numbers should be insignificant (NHS IC March 2012a).

The figures show an average annual increase, for the period 2000-2010, in FTE of 5.1% for consultants O&G FTE; 12% for registrars O&G FTE and 1.6% for midwives FTE (including bank staff).

Table 4.5-1 HC & FTE for Selected Staff Groups in Maternity Services, England

	September 2010	% increase since September 2009	% increase since September 2000	Average annual increase since 2000
Consultant O&G				
Head count	1,789	7.7% (128)	56.1% (643)	4.6%
FTE	1,725	7.8% (125)	64.5% (676)	5.1%
Registrars O&G				
Head count	2,888	5.1% (141)	207.6% (1,949)	11.9%
FTE	2,819	2.6% (72)	209.7% (1,909)	12%
Midwives (including bank)				
Head count	26,825	1.3% (341)	18.8% (4,253)	1.7%
FTE	20,790	2.7% (554)	17.7% (3,128)	1.6%
Midwives (excluding bank)				
Head count	24,938	2.4% (592)	19.5% (4,076)	1.8%
FTE	20,126	3.2% (629)	17.5% (2,996)	1.6%

Source: NHS IC September 2010 Census, p.23/24
https://catalogue.ic.nhs.uk/publications/workforce/numbers/nhs-staf-non-medi-2000-2010/nhs-staf-non-medi-cens-bull-2000-2010-rep.pdf

Table 4.5-2 below presents September 2010, Non-Medical Workforce Census (headcount and FTE) in maternity services in England, provided from the NHS IC which were used in the models. According to the data, there were 20,096 FTE midwives in the NHS maternity services in England in September 2010 and 3,707 FTE healthcare assistants and support workers.

Table 4.5-2 NHS Hospital and Community Health Services: Qualified Nursing, Midwifery & Support to Doctors and Nursing Staff in Maternity Services

England as at September 2010

	Maternity Services Including Bank Staff	
	Head Count	FTE
All nursing, midwifery & support staff	41,720	33,122.5
Nurse consultant	69	64.1
Modern matron	425	409.4
Manager	431	399.4
Registered nurse - Children	906	763.0
Registered midwife	25,451	20,095.8
Other 1st level[1]	5,729	4,806.7
Other 2nd level[1]	134	115.7
Support to doctors & nursing staff		
Nursery nurse	622	474.3
Nursing assistant/auxiliary	3,044	2,287.1
Healthcare assistant	3,551	2,746.1
Support worker	1,358	960.9

Source: The NHS Information Centre for Health and Social care

1. Other 1st and 2nd level include staff coded as Community Psychiatric Nurses (CPN) and Community Learning Disabilities Nurses (CLDN) with a specific recordable community qualification

Table 4.5-3 presents the medical staff groups (headcount and FTE) in Obstetrics and Gynaecology from September 2010 Medical Workforce Census in England, provided from the NHS IC which were used in the models. According to the figures, there were 1,724.5 FTE consultants O&G in the NHS in England in September 2010.

Table 4.5-3 Medical Staff in O&G by Grade, FTE and Head Count

England as at September 2010

	Obstetrics & Gynaecology	
	Head Count	FTE
Total	5,712	5,382.5
Consultant (including Director of Public Health)	1,820	1,724.5
Associate Specialist	163	134.3
Specialty Doctor	219	183.2
Staff Grade	64	54.2
Registrar Group	2,906	2,819.3
Senior House Officer	56	54.1
Foundation Year 2	297	295.6
House Officer and Foundation Year 1	100	99.6
Other Doctors in Training	-	-
Hospital Practitioner/ Clinical Assistant	75	12.9
Other Staff	12	4.7

Source: The NHS Information Centre, Medical and Dental Workforce Census, 2011

The main purpose of trust level workforce data was to create a quantitative measure for each staff group in each trust by first grouping some medical and non-medical staff groups to create four distinctive groups of consultants, doctors, midwives and HCA; adding the FTE for each staff group in each trust and dividing these by the ONS total number of births in each trust (so a staff group FTE per birth ratio for each trust was created). These were consultants FTE/birth; doctors FTE/birth; midwives FTE/birth and healthcare assistants (HCA) FTE/birth) for each NHS trust. These ratios were standardized at trust level and used in the models. The doctors group comprised of registrars, associate specialists and staff grade as well as trainees, i.e. house officers, senior house officers, specialty doctors and foundation year 2. These were all combined as 6 trusts had missing data on trainees. Healthcare assistants FTE

group comprised of HCA FTE and support workers FTE from the non-medical workforce file. Midwives comprised of all registered midwives.

Staff datasets (medical and non-medical) were subsequently merged with trust characteristics data (London trust/Foundation trust/Teaching trust) and trusts configuration data files (obstetric unit/alongside midwifery unit/free-standing midwifery unit) and together matched to HES patient-level data by using unique trusts identifiers.

Detailed staff groups and definitions of medical and non-medical staff in Maternity could be found in APPENDIX II.

4.6 DATA ON TRUSTS CHARACTERISTICS

Trust level characteristics included teaching status, London trust, foundation trusts status. The configuration of maternity service within trusts into Obstetric Unit (OU)/Alongside Midwifery Unit (AMU)/Free-standing Midwifery Unit (FMU) was adopted from the *Birthplace* in England (2011) study (See Chapter 2: Organisation of maternity services for definitions). This captured the choice of place of birth offered to women by their local NHS Trust. The non-foundation/foundation trusts divide was used to distinguish foundation trusts which are part of NHS but had more financial autonomy from the government. The foundation status was granted to top-performing hospitals. London was also distinguished from the rest of England because of high concentration of trusts and births in the capital (16% of all trusts providing maternity services in England were in London) and for having a more diverse population.

Table 4.6-1 shows that a third of the 143 trusts in the study were teaching trusts, more than half were foundation trusts, 26 (18.2%) of the trusts were in London; nearly half of the trusts (42%) had only an obstetric unit and only 14 percent had all the three types of units available to women namely an obstetric unit (OU), alongside midwifery unit (AMU) and freestanding midwifery unit (FMU).

Table 4.6-1: Characteristics of NHS trusts with maternity services

	Frequency	Percent
Teaching trust		
Not a Teaching trust	99	69.2
Teaching trust	44	30.8
Total	143	100
London trust		
Not a London trust	117	81.8
London trust	26	18.2
Total	143	100
Foundation trust status at September 2010		
Not a Foundation trust	69	48.3
Foundation trust	74	51.7
Total	143	100
Trust configuration		
OU only	60	42
OU and AMU	45	31.5
OU and FMU	18	12.6
OU and AMU and FMU	20	14
Total	143	100

**OU – Obstetric Unit; AMU – Alongside Midwifery Unit; FMU – Freestanding Midwifery Unit
Sources: DR Foster (Teaching trusts); MONITOR Web page (Foundation trusts status); BirthChoiceUK – Trusts configuration, OU/AMU/FMU; LSA (Local Supervising Authority) – London trusts from LSA website, LSA Annual Report to the NMC, 1st April 2011-31st March 2012, p.28. Imputed information on 3 London trusts RYJ (Imperial College Healthcare NHS Trusts); RYQ (South London Healthcare NHS Trusts); RYR (Western Sussex Hospitals NHS Trust).

4.6.1 MERGING OF TRUSTS

A few trusts were merged and pseudo trust identifiers were created and used when matching the different dataset. The rationale for combining some trusts or trusts and PCTs was that a trust or a PCT will have only midwives with links to an obstetric unit in another trust. Some deliveries took place in NHS establishments without maternity units, which were combined. Staffing data had to be added within the merged trusts/PCTs before being matched to HES. Table 4.6-2 below shows the amalgamation of the mentioned trusts/PCTs.

Table 4.6-2 Grouped Trusts/PCTs

Trust	Trust	Pseudo Trust Identifier
RJF (Queen's Hospital, Burton on Trent)	5PK (South Staffordshire PCT)	RJF/5PK
RH8 (Royal Devon and Exeter NHS Foundation Trust)	5QQ (Devon PCT)	RH8/5QQ
RTE (Gloucestershire Hospitals NHS Foundation Trust)	5QH (Gloucestershire PCT)	RTE/5QH
RD3 (Poole Hospital NHS Foundation Trust)	RDZ (The Royal Bournemouth and Christchurch Hospitals NHS Foundation Trust)	RD3/RDZ
RA7 (University Hospital Bristol NHS Foundation Trust)	RA3 (Weston Area Health NHS Trust)	RA7/RA2
All non-maternity trusts		OTHER

4.7 POPULATION USED IN THIS THESIS

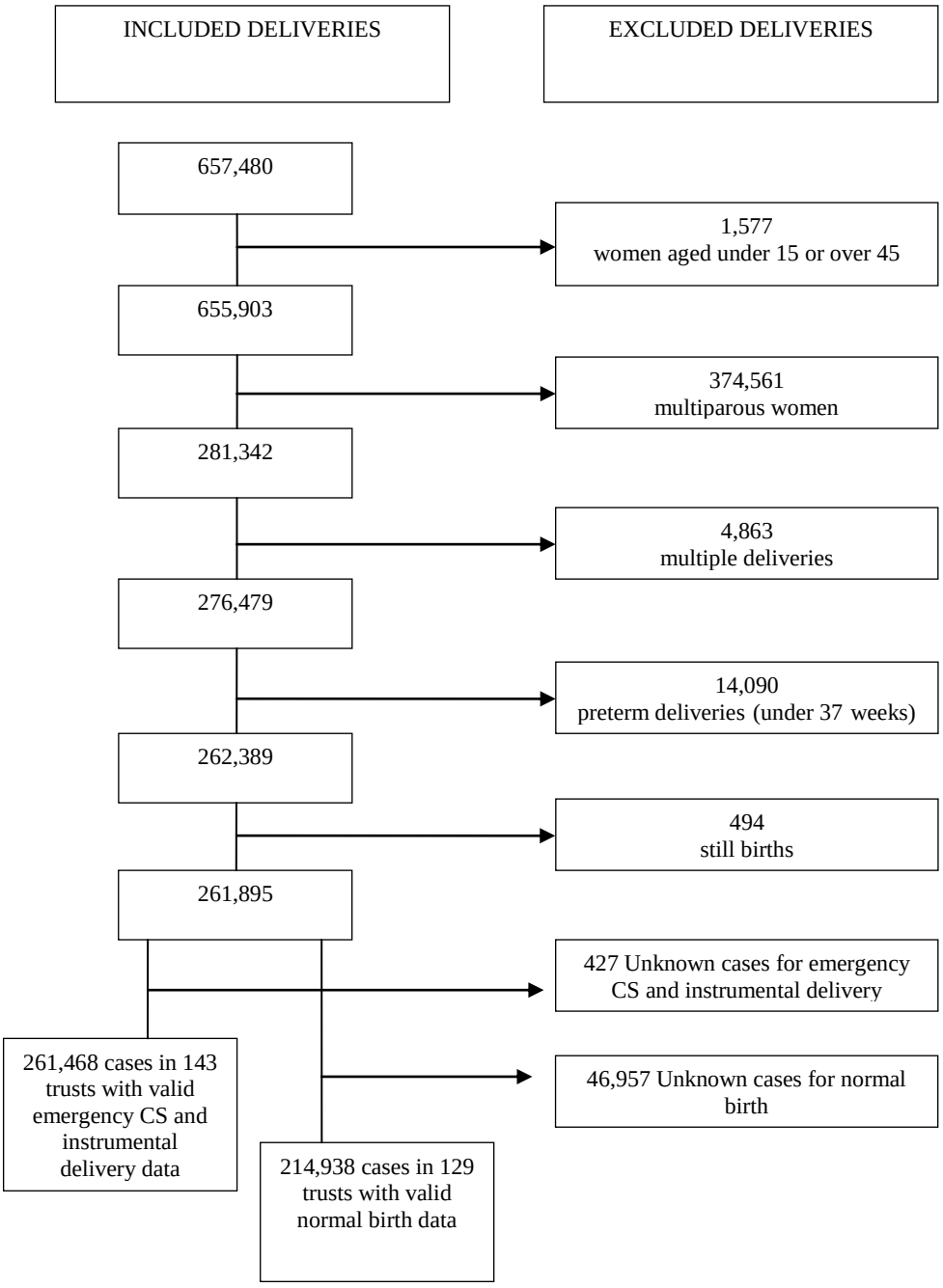
Maternity workforce data and trust characteristics were matched to HES patient level data using unique NHS trusts identifiers. After matching the datasets, the complete data consisted of 657,480 women in 143 NHS trusts with maternity services in England for the period April 2010/March 2011, while the trust level maternity staff and organisational data was for 143 NHS trusts as at September 2010.

In the HES population of 657,480 women in 143 trusts, there was no missing data for parity, but there were 368 missing cases for multiple births; 110,411 cases for gestational age; 73,785 cases for live/still births; 840 missing cases for emergency CS and instrumental delivery; 108,740 cases for normal birth and 1,868 missing cases for maternal age.

From the 657,480 women in the 143 trusts, selection was first made for women aged 15-44 years. 1,577 cases were dropped (for women under 15 and over 45 years) and 655,903 remained. The next steps were to select nulliparous women (281,342 cases left); singleton births (276,479 cases left); term births (>=37 weeks, 262,389 cases left); and live births (261,895 cases left). Figure 5 illustrates the data flow.

The missing data on maternal age, multiple births, gestation and live/still births were not discarded at the sample selection stage. It was assumed that all of the missing data were for singleton, live births. This of course may have introduced bias in the sample. The missing data for emergency CS, instrumental delivery and normal birth were dropped, as these were the dependent variables in the models. Consequently two sub-samples were created - one with all valid emergency and instrumental delivery cases (261,468 cases in 143 trusts) and another with all valid normal birth cases (214,938 in 129 trusts). When modelling normal birth, it also became apparent that one trust from the normal birth sub-sample contained only one case and another one contained overall 17 cases of which 10 were normal births. These two trusts were excluded in the multilevel model because of very small cell counts (the remaining sample in the normal birth model was therefore 214,920 deliveries in 127 trusts).

Figure 5: Data Flow



In the whole of HES 2010/11, data on gestational age, birth weight and live/still births, though improving in quality in recent years, included many missing cases (16.8%, 10.2% and 11.2% respectively). At trust level, few trusts had 100% missing data on these same variables:

- Gestational age (10 trusts with 100% missing and 16 trusts with >20% missing data)
- Infant birth weight (7 trusts with 100% missing and 11 trusts with >20% missing data)
- Ethnicity (2 trusts with >20% missing data)
- IMD (1 trust with >20% missing data)
- Live/still births (6 trusts with 100% missing data and 10 trusts with >20% missing data).

Data (in the cleaned HES) were less likely to be missing on maternal age (0.3%), ethnicity (4.5%), deprivation (0.8%), multiple births (0.1%), and delivery method (0.1%). The information on parity and NICE clinical risk was complete. Normal birth was a composite measure and because of its components was not available in 16.5% of the cases.

4.8 SELECTED INDICATORS/OUTCOMES

4.8.1 BIRTH MODE

Birth mode is recorded in HES maternity tail under the term delivery method (DELMETH). In addition the HES record of the mother contains OPCS delivery codes. Overall OPCS codes take precedence over maternity tail DELMETH variable.

The OPCS codes for each woman were searched for one of R17-R25 codes. These were mapped into a variable DELMETH_OPCS (see Table 4.8-1 below). The DELMETH_OPCS codes mirror those of DELMETH. In order to maximize all the available information, a third variable - best estimate of method of delivery (DELMETH_BEST) - was derived by using both DELMETH_OPCS and

DELMETH. When the DELMETH_OPCS was known (codes 0-8), DELMETH_BEST took the value of DELMETH_OPCS; when DELMETH_OPCS was unknown DELMETH_BEST took the value of DELMETH (if known and between 0-8). If both DELMETH_OPCS and DELMETH were unknown then DELMETH_BEST took the value of DELMETH_OPCS.

Table 4.8-1 Delivery Method (DELMETH) from HES maternity tail

DELMETH Maternity Tail HES	**DELMETH_OPCS**	**OPCS Code**
0 = Spontaneous vertex (normal vaginal delivery, occipitoanterior)	0	~ R24*
1 = Spontaneous other cephalic (cephalic vaginal delivery with abnormal presentation of head at delivery, without instruments, with or without manipulation)	1	~ R23*
2 = Low forceps, not breech, including forceps delivery not otherwise specified (forceps, low application, without manipulation)	2	= (R215 V R218 V R219)
3 = Other forceps, not breech, including high forceps and mid forceps (forceps with manipulation)	3	= (R211 V R212 V R213 V R214)
4 = Ventouse, vacuum extraction	4	~ R22*
5 = Breech, including partial breech extraction (spontaneous delivery assisted or unspecified)	5	~ R20*
6 = Breech extraction not otherwise specified, including total breech extraction and version with breech extraction	6	~ R19*
7 = Elective caesarean section (caesarean section before or at the onset of labour)	7	~ R17*
8 = Emergency caesarean section	8	~ R18*
9 = Other than those specified above, including destructive operation to facilitate delivery, and other surgical or instrumental delivery, for example, application of weight to leg in breech delivery	9	~ R25*
null = Not applicable (from 1990-91 to 1994-95)	X	Otherwise
x = Not known (from 1996-97 onwards)		

Source: Rod Gibson Note: v means "or"

DELMETH_BEST was an important derived variable because all the outcomes investigated were either directly based on it (emergency CS and instrumental delivery); or derived from it in the case of normal birth. Normal birth was a composite indicator and was defined as birth without: induction, episiotomy, general and/or regional anaesthetic and not being instrumental or caesarean. As mentioned earlier the study by Knight et al. (2013), and cited by the RCOG report (2013) reported that the mode of delivery was well recorded in HES, with strong levels of

internal agreement between OPCS delivery codes in HES and DELMETH in the maternity tail.

4.8.2 NORMAL BIRTH

Normal birth as used in the models was derived from:

- Method of delivery (DELMETH_BEST=0, 1 and 5, i.e. women with spontaneous vertex, spontaneous other cephalic delivery, breech including partial breech extraction) and;
- Method of onset of labour (DELONSET=1 for spontaneous onset of labour, excluding values 3,4,5 for medical and/or surgical induction and value 2 for any CS before onset of labour) and;
- no episiotomy[93] and;
- no induction[94]; and
- no mention of either pre- or post-delivery use of general or regional anaesthetic (in DELPREAN, DELPOSAN or OPCS)[95]
- with a denominator made of all usable[96] records from Method of delivery (DELMETH_BEST), Method of onset of labour and known anaesthesia.

---

[93] Episiotomy is defined as incision at the vaginal opening to facilitate birth and was extracted from any record with an OPCS codes R27.1; R27.8; R27.9 indicating an episiotomy with a denominator made of all vaginal deliveries coded in Method of delivery (Delmeth_Best from 0 to 6).

[94] Induction is derived from Method of onset of labour (Delonset=3, 4 or 5, i.e medical and/or surgical induction with a denominator made of all usable records for which method of onset is known).

[95] Anaesthetics used is derived from DELPREAN (anaesthetic used before delivery), DELPOSAN (anaesthetic used post delivery) and OPCS codes Y80, Y81, Y82, Y84 and Y90 (which cannot distinguish between the two timings of anaesthetic given).

[96] 'Usable Records' meant that all non-delivery events, duplicate records and home births were excluded.

This definition of normal birth (which excludes both pre- and post-delivery anaesthetic) is stricter compared to the normal birth definition from the Maternity Care Working Party (MCWP 2007). The definition of MCWP[97] 2007 for normal birth (spontaneous delivery, spontaneous onset of labour, no pre-delivery anaesthesia and no episiotomy) included women with post-delivery anaesthetic. Though different from the MCWP definition, this derived normal birth indicator which was used in the modelling, benefits from the improved accuracy of anaesthetic information derived from the OPCS codes as well as DELPREAN and DELPOSAN. This increased the number of women in the normal birth denominator. Also very few women who had a normal birth according to the MCWP definition were expected to develop complications requiring a post-delivery anaesthetic.

4.8.3 EMERGENCY CAESAREAN SECTION

Emergency CS was taken from Method of delivery (DELMETH_BEST=8) with a denominator made of all usable records for which Method of delivery was known.

4.8.4 INSTRUMENTAL DELIVERY

Instrumental delivery was taken from Method of delivery ((DELMETH_BEST = 2, 3 and 4; i.e. 2 = low forceps, not breech, including forceps delivery not otherwise specified (forceps, low application, without manipulation); 3 = other forceps, not breech, including high forceps and mid forceps (forceps with manipulation); and 4 = ventouse, vacuum extraction)) with a denominator made of all usable records for which Method of delivery was known.

---

[97] MCWP definition of normal birth was derived from DELONSET (1=for spontaneous onset of labour); DELMETH_BEST (0, 1 and 5, i.e. women with spontaneous vertex, spontaneous other cephalic delivery, breech including partial breech extraction); no episiotomy (as above) and DELPREAN (if not equal to 1,2,3,4,5,6, i.e. general, epidural or caudal and spinal anaesthetic and any combination of these with a denominator made of all usable records from Method of delivery (DELMETH_BEST), Method of onset of labour and known anaesthesia.

4.9 BACKGROUND VARIABLES

The models in this thesis used variables which were considered important in predicting the three outcomes – emergency CS, instrumental delivery and normal birth, i.e. *"….decisions about inclusion of covariates should be based on appropriate clinical knowledge and judgement in conjunction with statistical results"* (Sinha et al. 2012:4).

4.9.1 MATERNAL AGE

Studies have shown that maternal age greater than 35 is a risk factor for a range of co-morbidities (CEMACH 2005; CMACE 2010). These include: gestational diabetes, pregnancy induced hypertension, pre-eclampsia and type 2 diabetes (see also Teenage mothers and mothers aged >40 in Chapter 2). There is also an increased risk of caesarean section for older women (Paranjothy et al. 2005), while women under the age of 20 experienced higher rates of stillbirths and had higher rates of perinatal and neonatal death when compared to women aged 20-34 (CEMACH 2009).

Given that clinical risk increases with maternal age, maternal age was considered important predictor for the three outcomes. Maternal age was grouped into 6 categories: 15-19, 20-24, 25-29, 30-34, 35-39, 40-44 in the models. The groupings will allow for modelling non-linear relationship between maternal age and outcome variables. These were also the groupings used in the RCOG (2013) report; in Paranjothy et al. (2005); Bragg et al. 2010 and Essex et al. (2013).

4.9.2 WOMAN'S ETHNICITY

Results from a National cohort study of ethnic variation in severe maternal morbidity using the UK Obstetric Surveillance System (UKOSS) from 2006 (Knight et al.

2009), showed that Black African women and black Caribbean women had the highest risk of severe maternal morbidity[98] compared with white women. Non-white women had higher risk of severe maternal morbidity even after adjustment for differences in age, socioeconomic and smoking status, body mass index, and parity.

A national cohort study from UK in 2009 of women who gave birth in UK during 2 months in 2009 (CMACE 2010) found that Black and Minority Ethnic (BME) group with maternal obesity (BMI ≥35) were 3.5 times more likely to have type 2 diabetes and 1.6 times more likely to have gestational diabetes than White women with a BMI ≥35. After controlling for diabetes, BME women were still more likely to have a caesarean section, have a pre-term baby (before 37 weeks' gestation) and longer hospital stay (after both vaginal deliveries and caesarean sections). Women with type 2 diabetes were also more likely to be from a Black, Asian or Other Ethnic minority group, than women with type 1 diabetes (CEMACH 2005).

The cleaned ethnic categories of women in HES were aggregated under the following 7 groups:

- White ('British White'; 'Irish White'; 'Any other White background' ; 'White British Irish Other White old');
- Asian ('Indian Asian or Asian British'; 'Pakistani Asian or Asian British'; 'Bangladeshi Asian or Asian British'; 'Any other Asian background'; 'Indian old'; 'Pakistani old'; 'Bangladeshi old');
- Black Afro-Caribbean ('Caribbean Black or Black British'; 'African Black or Black British'; 'Any other Black background'; 'Black Caribbean old'; 'Black African old'; 'Black Other old');
- Mixed ('White and Black Caribbean Mixed'; 'White and Black African Mixed'; 'White and Asian Mixed'; 'Any other Mixed background');

[98] Severe maternal morbidity – women diagnosed with acute fatty liver of pregnancy, amniotic fluid embolism, antenatal pulmonary embolism, eclampsia and peripartum hysterectomy.

- Chinese ('Chinese other ethnic group'; 'Chinese old');
- Other ('Any other ethnic group'; 'Any other ethnic group old');
- Not known/not stated/not given ('Not given old'; 'Not known old and new'; 'Not stated').

At the modelling stage Mixed, Chinese and Other categories were aggregated due to small numbers and a consideration for degrees of freedom in the modelling stage.

4.9.3 INDEX OF MULTIPLE DEPRIVATION (IMD)

38 indicators from 7 weighted domains are used to create the Index of Multiple Deprivation 2010 overall ranking (Department for Communities and Local Government 2011). These domains measure area socio-economic deprivation, not individual. The seven domains and their respective weights are:

- Income (22.5%)
- Employment (22.5%)
- Health Deprivation and Disability (13.5%)
- Education, Skills and Training (13.5%)
- Barriers to Housing and Services (9.3%)
- Crime (9.3%)
- Living Environment (9.3%)

Studies have shown that social deprivation was associated with maternal obesity (Heslehurst et al. 2007); a higher percent of obese pregnant women in England (34% with BMI ≥35) were in the most deprived quintile (based on IMD) compared to 27.6% for all maternities (CMACE 2010); and that maternal social deprivations (based on postcode of residence) was associated with poor pregnancy outcomes (CEMACH 2005).

IMD overall measure was first re-coded into quintiles of deprivation in the HES working file for all deliveries (see descriptive statistics in Chapter V). As the interest

was in comparing birth outcomes for women from the most deprived and the least deprived areas in England, least deprived was used as a reference category in all the models and the initial quintiles coding was reversed, such that 1= Least Deprived and 5=Most Deprived. In the 2010/11 HES data with 657,480 deliveries (used in the models, see Chapter V), there were 0.8% missing or unknown records for area of deprivation (IMD) allocated to a woman.

4.9.4 BIRTH WEIGHT

Birth weight was defined as infant birth weight in grams immediately after birth (HES definition). Infant birth weight (BW) was re-coded into three categories:

- BW less than 2500g
- BW more than 2500g (incl) and less than 4500g (incl)
- BW more than 4500g

According to HES 2010/11 online[99], Asian and Asian British women (8.7%), followed by Black and Black British (6.8%) had the highest percentages of infants born weighing less than 2500g, while the white ethnic group had the highest percent (13%; while Asian had only 5%) of infants weighing more than 4000g. For live singleton births in HES 2010/11 online, 82.7% of infants were born weighing between 2500g and 3999g; 5.5% were under 2499g and 11.8% were above 4000g. It has also been reported that birth weight was higher now compared to 20 years ago and that demographic changes (including birth weight) may have contributed to an increase in CS rates; it was suggested that comparison of CS rates should control for at least for maternal age, birth weight and parity (Parish et al. 1994 & Kirsop et al. 1992 cited in Paranjothy et al. 2005).

[99] http://www.hscic.gov.uk/pubs/maternity1011, Table 28 and Table 29.

4.9.5 GESTATIONAL AGE

Gestational age was defined as: *"The age of the fetus or newborn calculated from the number of completed weeks since the first day of the woman's last menstrual period"* (NICE 2008).

According to the epidemiological evidence reviewed in NICE (2008) induction of labour guideline, there were associated risks for the infant and increased risk of interventions for women (such as CS) if pregnancy progresses beyond 40 weeks gestation. These risks were however small and the systematic review data used in the guideline indicated that these were not reduced by induction of labour (though the studies did not have sufficient statistical power to address this question). NICE (2008) guideline recommended that certain benefits from induction, had to be balanced with risks and complications but that induction of labour for prevention of prolonged pregnancy should be offered from 41+0 weeks onwards (NICE 2008, Induction of Labour CG 70).

Gestational age was grouped into two categories in the analyses: 37-41 weeks of gestation; greater than 42 weeks of gestation.

4.9.6 SMOKING

Smoking[100] was not used in the models because of quality issues. The reported rates by trusts ranged from 0% to 25%. It was difficult to distinguish underreporting from genuine low rates.

[100] ICD codes in HES indicating smoking were F16, Z72.0, Z71.6 and P04.2

4.9.7 DEFINITION OF NICE (2007) CLINICAL RISK COMPOSITE MEASURE

The rationale for creating a composite measure of risk was that women with certain clinically recognised medical conditions unrelated or related to previous or present pregnancies have an increased risk of complications or adverse outcomes. Therefore it was hypothesised that the outcomes for these women and their babies differ from women considered to be at lower risk. In addition different trusts may have different proportions of these women and thus their maternity indicators may differ. To account for these differences a composite measure of clinical risk was created.

Differentiating between higher and lower risk pregnancies in this study was based on the methodology of the *Birthplace* in England (2011) study. The study compared safety of birth in different settings for women judged to be at 'low risk' of complications at labour onset. *Birthplace* definition of risk was based on NICE (2007) Clinical Guideline 55 Intrapartum Care to define women at 'low[101] risk'. The conditions listed in the NICE 2007 guideline were either pre-existing or have developed during the current pregnancy, but they were present at the end of pregnancy or onset of labour (aiming for timely referral of women to care in obstetric unit). NICE (2007) contains four tables, (see Appendix V); two relate to medical conditions (i.e. confirmed cardiac disease) and other factors (i.e. eclampsia), both indicating increased risk and suggesting planned birth at an obstetric unit; the other two tables relate to medical conditions (i.e. cardiac disease without intrapartum implications) and other factors (i.e. previous pregnancy complication such as third or fourth degree perineal trauma) which require individual assessment when choosing place of birth. These conditions include among others diabetes and gestational diabetes, hypertension, eclampsia, epilepsy, renal disease, cardiac disease, liver

[101] Low Risk Women were classified as such if they were not known to have any of the medical conditions or situations listed in the NICE Intrapartum Care guideline (2007) immediately prior to the onset of labour. If any of these medical conditions were recorded, a woman (and her baby) were seen as having an increased risk during and shortly after labour and care in an obstetric unit where the risk was expected to be reduced was recommended.

disease, asthma, anaemia, cystic fibrosis, infections, breech presentation, shoulder dystocia, previous CS, previous post partum haemorrhage, placenta praevia and placenta abruption, previous unexplained still births, etc.

Three categories of risk status were initially derived from NICE (2007) – NICE (low); NICE (individual assessment) and NICE (increased). These three categories were created by extensively mapping and verifying the medical conditions in the NICE (2007) Clinical Guideline 55 Intrapartum care against the relevant 4 digits ICD-10 codes (around 12,000 of them), some OPCS codes and HES Data Dictionary data items for each delivery record in HES.

Women were classified as low risk if, immediately prior to the onset of labour, they were not known to have any of the medical conditions or situations listed in the NICE Intrapartum Care guidelines that result in increased risk for the woman or baby during or shortly after labour, where care in an obstetric unit would be expected to reduce this risk.

In this research women were categorised as having a 'higher risk' pregnancies if they had:

- pre-existing medical condition or;
- complicated previous obstetric history or;
- conditions which develop during pregnancy.

Using information contained in women's HES records, women were retrospectively and systematically allocated to lower and higher risk status at the end of pregnancy. Because some of the conditions related to events arising from previous pregnancy and delivery, 10 years of previous linked in-patient records of women were searched for such conditions. Pre-existing conditions not related to pregnancy (e.g. cardiac disease) were also searched for in any previous linked in-patient record. A few additional conditions not listed in the guidelines were also mapped; these were also seen to contribute to an increased risk.

Induction was excluded from the list of NICE (2007) conditions, even though induction carries increased risk for women who are therefore expected to give birth in an obstetric unit according to NICE (2007). Induction is a procedure, not diagnosis and the risk of a woman at the end of pregnancy should be determined by diagnosis of a condition rather than the decision to induce.

Certain limitations of the methodology for assigning risk were acknowledged by the authors (see Notes on p. 155-56). It was recognised that each condition may have a 'spectrum of risk', and be associated with different adverse outcomes. Not every code was allocated to a NICE risk category, due to the large number of ICD-10 codes therefore the proportion of higher risk women may have been underestimated. However it was unlikely that these unallocated codes would relate to NICE (2007) conditions. There was not always a direct match for conditions and in some cases it was not possible to determine whether a diagnosis was related to the antenatal, intrapartum or postpartum period. In these situations careful judgements were made. Assignment of risk was also affected by incomplete or inaccurate coding and trusts with poor coding may present with a higher proportion of lower risk women. Despite these limitations the authors were confident that the risk assignment tool distinguishes well between the groups of women. However the tool was recommended for use only in retrospective analysis rather than to prospectively categorise women in a clinical context.

In the models of this thesis NICE (increased risk) and NICE (individual assessment) were combined and used as a single category "NICE High Risk", because of a concern for small proportion of women belonging to the individual assessment category.

4.10 RISK ADJUSTMENT

Patient characteristics and NICE clinical risk available from HES 2010/11 were used for risk adjustment in the models to control for the case mix differences across trusts.

Risk adjustment is a technique used in observational studies to identify and control for potential confounders (socio-demographic and clinical risk factors of the mother and the fetus) when comparing outcomes between providers. These factors are not expected to be homogeneously distributed among the populations in different trusts (Iezzoni 1997; Stivanello 2013).

The following is an explanation of why risk adjustment was applied in the models used in this thesis.

Models are usually risk adjusted in order to compare outcomes between hospitals (Fantini et al. 2006). Baghurst (2012) in a study of all (65,598) singleton vaginal births in 18 public maternity hospitals in South Australia aimed to understand the time trends (2002-2008) in incidence of $3^{rd}/4^{th}$ degree maternal perineal tears and established that using risk factors which were partly subjective or related to clinical practices could "*mask the underlying reasons for differences in the outcome under consideration"* (Baghurst 2012:4).

Fantini et al. (2006) in a study of risk adjustment for inter-hospital comparison of primary caesarean section rates warned that a large number of potential confounders were being used in studies of caesarean rates, including socio-demographic, maternal and fetal clinical conditions and organizational factors, even though no consensus existed on the actual confounders. In addition the authors warned that using many confounders could be problematic in relation to data completeness, accuracy and reliability which could reduce the precision of the adjusted measures (Fantini et al. 2006). Peaceman et al. (2002) in their study of risk-adjustment of caesarean delivery rates, argued in the case of caesarean section that diagnoses such as fetal distress, failed induction, disproportion, etc. could be related to clinical decision making and

thus should be excluded from the risk-adjustment model, because they represent post-hoc justification for performing a caesarean section rather than being risk factors; consequently they appear to be strong predictors of caesarean deliveries.

Stivanello et al. (2013) in an observational study using administrative data investigated the issue of how many variables were needed in risk adjustment for caesarean delivery for inter-hospital comparison. The study showed that the models with more variables explained more variation in the caesarean delivery rates but that simpler adjustment worked as well as complex ones; and that when most important factors for caesarean delivery were omitted, the predictive power of the models was poorer but the observed to expected ratios of CS rates were similar across models.

Factors such as infant birth weight, gestational age and parity are not absolute indicators of CS delivery (Fantini et al. 2006) but were used by other researchers (Fantini et al. 2006; Peaceman et al. 2002; Bragg et al. 2010; Bailit et al. 2006) because they were considered important and frequent predictors of CS deliveries (Stivanello et al. 2013). Gestational age and birth weight though are likely to be correlated.

Bailit et al. (2008) in a study of the role of race in primary caesarean rate case-mix adjustment used 2003 California birth certificates data and found that race/ethnicity could be left out from risk adjustment models for primary CS, on the basis of having a small impact on the predictive quality of the models in their study. If left in they could not *'explain away'* the large differences in outcomes in a context of discrimination. The study discussed the controversial nature of including race/ethnicity for risk adjustment. The question was whether race/ethnicity differences were based in economics, biology or discrimination. In the first two instances there was a legitimate reason to include them in the risk-adjustment; in the third instance they were not appropriate for inclusion because of the possibility of masking important social issues (Bailit et al. 2008).

In UK a national cohort study by Knight et al. (2009b) aimed to establish the ethnic differences in severe maternal morbidity (acute fatty liver of pregnancy, amniotic fluid embolism, antenatal pulmonary embolism, eclampsia or peripartum hysterectomy) for 2005-06, using information from UKOSS[102]. The estimated incidence was 89 per 100,000 maternities. The study concluded that severe maternal morbidity was more common among non-white women compared to white women in UK and particularly for black African and Caribbean ethnic groups. The pattern was similar to differences in maternal death rates. The risk remained after adjusting for age, socioeconomic status, smoking, BMI and parity. The authors speculated that the differences were possibly due to pre-existing medical conditions and/or factors related to care during pregnancy, labour and birth (particularly in relation to poor and late access to antenatal care). Regional studies from the United States have shown that black and other non-white pregnant women had higher rates of pre-existing hypertension and diabetes (Rosenberg et al. 2005).

Clinical diagnoses such as shoulder dystocia, fetal distress, failure to progress and induction of labour were not considered in the risk adjustment for the three models in the current study, as their assessment could be influenced by practice style rather than reflecting patient characteristics or clinical risk. NICE guidelines on induction of labour (2008) for example highlight that there is no general agreement on criteria for failed induction. The recommended management after failed induction is for another attempt to induce labour or to opt for caesarean section. Although there is a risk of caesarean delivery after a failed induction, induction was not included in the model for emergency CS, as it was seen to reflect clinical decision making and not patient characteristics or clinical risk. The same argument applies to use of continuous electronic fetal monitoring during labour, despite evidence suggesting that for low-risk women it was associated with an increase in emergency caesarean section (Alfurevic et al. 2006).

---

[102] UKOSS – the UK Obstetric Surveillance System. The system allowed for routine study of severe maternal morbidity on national level.

In the current study a sample of all mothers aged 15-44, who were nulliparous, at term, with a singleton, live birth was selected and analysed. Additionally, the three models for emergency CS, instrumental and normal birth, all controlled for maternal age, ethnicity, deprivation (IMD), clinical composite risk (NICE 2007), gestational age and birth weight to allow for comparison between models and to isolate variations in outcomes between trusts due to staffing levels and organisational factors. The sample was homogeneous for parity, singleton/live births and at term deliveries (gestational age >37 weeks) but not for birth weight and clinical risk. The same socio-demographic and clinical predictors were used for the three outcomes to allow for comparison across models. Data on lifestyle factors such as obesity, smoking and alcohol consumption were either of poor quality or not available in HES.

4.11 STATISTICAL ANALYSES

HES 2010/11 and NHS IC workforce datasets were matched before the modelling stage, i.e. women (Level 1) were nested within trusts (Level 2). Multilevel binary logistic models were employed because of the hierarchical nature of the data and the categorical nature of the outcomes (for example YES, did have an emergency CS=1 / NO, did not have an emergency CS=0). Multilevel binary logistic models were applied separately to each categorical outcome to model their relationship with maternity FTE staff/birth ratios and trust characteristics, by simultaneously controlling for maternal socio-demographic characteristics and clinical risk factors.

4.11.1 THE CONCEPT OF MULTILEVEL MODELLING

Health outcomes research tends to compare patient and organisational outcomes across institutions. Some of the data used in these studies is naturally hierarchical in structure (i.e. patients within hospitals). Multilevel models are analytic types of models that include variables measured at different levels of the hierarchy. By acknowledging the hierarchical structure, multilevel models allow for simultaneous modelling of the individual and group level predictors on the dependent variable of

interest. The conceptual and statistical advantages are that the variables are analysed at the level they were defined and measured (Cho 2003) with no requirement for aggregation or disaggregation of the data to fit a single level analysis (Hox 1995). For example, if trust level analysis is needed, there is no need to aggregate women characteristics (age, ethnicity) measured at individual level to a higher (trust) level (by calculating the overall rate or mean for each trust). Alternatively there is no need to include dummy variables for each trust (in this study 143) if a single level logistic regression at the level of the patient (women) is used.

Another major advantage of multilevel modelling approach is the ability to explore effects of group level (trust level) predictors, while accounting for the effect of unobserved group characteristics. They are particularly useful for exploring contextual effects (i.e. staffing, trust characteristics) as the standard errors of the group level coefficients can be inaccurate if a single-level model was used.

Several problems exist if the hierarchical structure of the data is ignored and a single level analysis applied. If the analysis is performed at a group (trust) level, a multivariate model could be used to deal with aggregated individual characteristics (mean maternal age) and group level variables (staffing). It is problematic to analyse the relationship between variables at group level and to draw conclusions at individual level. Making inferences in such way lead to the so called "ecological fallacy" (Robinson 1950), i.e. what is true at one level is not necessary true at another level. For example if the mean maternal age in each trust is significantly and positively associated with the overall emergency CS in trusts, this relationship does not allow any conclusion about the effect of age on the likelihood of emergency CS at individual level. When a variable is aggregated to group level, the interpretation of the variable does not refer to the individual anymore but to the group (Snijders and Bosker 1999). In the above example the mean maternal age relates to the population of the trust but its interpretation may be different from that of age at an individual level. Aggregating individual level data also eliminates the within-group (trust) variations and may inflate the estimates of the true relationship between variables (Kreft and De Leeuw 1998).

Alternatively, if the analysis is performed at the level of the individual with added dummy variables for the groups (trusts), it ignores the clustering of individuals within groups (patients within trusts), and violates the linear regression assumption of independence of observations, thus the precision of the estimates is affected. The standard errors of the predictor coefficients will be underestimated and inferences invalid, leading to increased risk of Type I error. The effects of the predictors may be misinterpreted in size but also in direction. Moreover trust level variables like staffing cannot be modelled.

Multilevel regression models are known in the literature as hierarchical linear models; generalized linear mixed models (GLMM); random-coefficient models/random-coefficient regression; mixed effect models; random effects models and growth models. "Multilevel models" is a favoured term in sociology, related to the idea that regression intercepts and slopes at the individual level may be modelled as random effects of a higher group level; "covariance components models" term is often used in statistics (meaning that the covariance can be decomposed into parts attributable to within-groups and between-groups effects); and in economics, the term "random coefficient regression models" is preferred (Garson 2013). These different names reflect methodological developments in several different disciplines over the last twenty years and the variations in methods and software programmes used in the different fields (Heck et al. 2012).

Among the first to develop the main concepts and methodological principals of multilevel data were Mason et al. (1983); Bryk and Raudenbush (1992) and Goldstein (1987, 1995). The methods were first applied by sociologists in educational research exploring school effects in respect to pupils' academic achievements. Interests in binary outcomes and hierarchical social structures led to the development of multilevel logit models (Wong and Mason 1985; Anderson and Aitkin 1985; Goldstein 1991).

The review of the multilevel modelling theory will proceed with comparison of ordinary regression and multilevel models for continuous response and statistical

presentation and issues with basic multilevel models for continuous response; models with categorical outcomes; single level logit models; multilevel logit models and generalised linear mixed models (GLMM); latent presentation model and variance partition coefficient (VPC).

4.11.2 ORDINARY REGRESSION VS MULTILEVEL MODELLING

Field (2009) in a textbook presented an intuitive interpretation of multilevel models. That is, ordinary regression technique assumes the data are organized at a single level, while multilevel methodology models hierarchical data. If analysis is performed at an individual (single) level only as done with ordinary regression, the existence of 'clustering' in the population is ignored, despite its prevalence in real data. One of the main assumptions when applying ordinary regression to a single level data is the independence of observations (i.e. patients). With hierarchical data, observations (i.e. patients within trusts) are not treated as independent anymore. For example one could expect that outcomes for patients as a result of a treatment within trusts to be similar compared to outcomes for patients from different trusts. This similarity is expected because patients within trusts are being treated by the same professionals, therefore the outcomes of their treatment could be similar (after controlling for individual socio-demographic differences and medical risk factors) and may be determined by the staff numbers, staff deployment, skills and experience, as well as the individual trust characteristics, including organization policies and more broadly the culture of the place.

Multilevel modelling methodology is applied to assess variations at trust and patient level simultaneously. For example, we can assess how much an outcome varies between trusts and how much it varies between patients within trusts simultaneously. Intra-class correlation coefficient (ICC) is used to measure the dependency between cases (patients), when the response variable is count or continuous in nature. The ICC measures the proportion of the overall variation in an outcome attributable to trusts, i.e. how similar are the outcomes within trusts. The higher the value, the more similar are the patients in trusts with respect to the outcome. If the trusts (their

staffing, organisation and policies) have a big influence on patients (on their treatment and outcome of treatment) within them, the variability within the trusts will be small and the health outcomes will be similar. For example a greater ICC will indicate that women within trusts are more likely to share common experiences. If trusts characteristics have little effect on patients, then the outcomes will vary a lot within trusts, which will make differences between trusts relatively small, i.e. ICC will be small. A note of caution, ICC interpreted in such ways is applicable to outcomes which are continuous or count in nature. The ICC is difficult to calculate for binary response outcomes and its interpretation is slightly different as explained further in this chapter (see Variance Partition Coefficient (VPC)).

With ordinary regression it is assumed that the parameters are fixed, i.e. the outcome, the predictors and the residuals all vary as a function of a particular case (observation/patient) in the data, while the parameters of the intercept and the slope are fixed. There are no consistent definitions of random and fixed effects in the literature. The following is taken from McMahon et al. (2006:2) *"Fixed effects are model components that assume no random variance or sampling error (e.g., group means, experimental conditions), and are constant across units of a given level; whereas random effects are model components that estimate population variance including sampling error (e.g., residuals, unobservable random quantities), and exhibit variation across units of a given level according to an error distribution. Multilevel models containing both fixed and random effects are referred to as mixed models"*.

Multilevel models allow parameters to vary (i.e. random effect). One can allow for random intercept, random slope or both and allow the relationship between predictors and an outcome to vary across trusts. This is achieved by permitting the intercept to vary across trusts, or the slopes to vary across trusts or both to vary across trusts. Technically this is done by adding a random component to the intercept, slope or both. One part of the resulting term then measures the intercept and the slope of the overall model fitted to the data and the other measures the variability of the intercepts and the slopes around that overall model. If we allow for

the intercept to vary we assume that the relationships between the predictors and an outcome are the same in all trusts, but that each trusts has a different intercept. If we allow for the slope to vary we assume that the relationships between the predictors and the outcome are different in each trust.

In other words, multilevel modelling approach helps to understand and measure how group membership introduces additional sources of variation in the data (Bliese 2012). The following is based on "Multilevel modelling in R" manual by Bliese (2012). Two additional variance terms distinguish the multilevel model for continuous response from an ordinary regression:

- Random Intercept Variance - measures the degree to which trusts differ in their mean value (intercept) on the outcome. A significant variance term means that trusts differ on the outcome, and allows one to include predictors which predict why some trusts have higher than average outcome values while others have lower than average outcome values. The trusts mean differences are predicted with trust level predictors (Level-2). These predictors have the same values within trusts but are different across trusts (i.e. staffing levels).
- Random Slope Variance – this term measures the extent to which slopes between the outcome and the predictors vary across trusts. Within a single regression this relationship is constant across trusts. Within multilevel analysis one can test if the slopes vary from one trust to another. If the slopes significantly vary then Level-2 predictors can be used to explain why the slope (relationship) between the outcome and a predictor is steeper (stronger) in some trusts than others (the variation is a function of trust differences).

A third variance term is common in both ordinary regression and multilevel models. It estimates within trust variations. This variance explains the degree to which an individual outcome differs from its predicted value within specific trust. Individual (Level-1) predictors are used to predict within trust variance term. These individual

level predictors (age, ethnicity, IMD, etc.) differ among the individuals within the same trust.

In the current study a two-level multilevel model was applied, i.e. women delivering (Level-1) within NHS trusts (Level-2) because of the availability of data at these two levels. One could consider three or four levels of nested data in a multilevel model, i.e. women within maternity units, units nested within maternity hospital sites, and hospital sites nested within NHS trusts with maternity services. Detailed data were not available to consider four levels model and even if it were available multilevel analysis may have not been feasible as one has to consider the number of observations at each level. The purpose of this research was to investigate the association between birth outcomes and maternity staffing levels and detailed maternity staff data from NHS IC was only available at trust level (not at unit or hospital level). Administrative data from HES could also not distinguish between women delivering at different maternity units (OU/AMU/FMU) or hospitals within trusts. In addition different datasets had to be matched using unique identifier, which was only available for NHS trusts. These limitations restricted the analyses to a two level multilevel model to include women giving birth (Level-1) within trusts (Level-2).

4.11.3 BASIC MULTILEVEL MODEL FOR CONTINUOUS RESPONSE VARIABLE

The basic generalized linear random intercept model for continuous response (e.g. a 2-Level random intercept multilevel model for continuous response) is presented in the form:

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + u_j + e_{ij}$$

$$u_j \sim N(0, \sigma_u^2) \; and \; e_{ij} \sim N(0, \sigma_e^2)$$

where y_{ij} is the response for individual i in group j; β_0 is the constant term 'intercept' and, β_1 is the coefficient of an individual level predictor x_{ij}; e_{ij} is the individual

level residual (Level-1 error term) and u_j is the group level residual (Level-2 error term). When no predictors are included, the model is called "unconditional" or "Null" model. What the "Null" model states is that the dependent variable is a function of a common intercept and two errors: the between group error term u_j and within group error term e_{ij}. It estimates how much variability there is in the mean y_{ij} (captured through the intercept variability) relative to the total variability. The error terms $e_{ij}\ and\ u_j$ are normally distributed with means 0 and variances σ_e^2 and σ_u^2 respectively.

The expectation (mean) of the response y_{ij} as a function of the predictor x_{ij} and u_j is expressed as:

$$E(y_{ij}) = \beta_0 + \beta_1 x_{ij} + u_j$$

The variance σ_u^2 of the intercept random effect, u_j, and the variance σ_e^2 of the Level-1 residual term, e_{ij}, are estimated and used to evaluate ρ, the Intra-class Correlation Coefficient (ICC). The group variance σ_u^2 measures how much each group's mean outcome (intercept) varies from the overall mean outcome (intercept), while the individual level variance σ_e^2 measures how much each individual's outcome differs from the group mean outcome. The ICC is a measure of the extent to which observations within a cluster are related and expressed as a ratio of the between-cluster variance to the total variance:

$$\rho = \frac{\sigma_u^2}{\sigma_e^2 + \sigma_u^2}$$

In multilevel linear models for continuous response, the ICC is also interpreted as the proportion of the total variance attributed to clustering. However, this standard ICC formulation is not valid in the case of multilevel models for binary response.

4.11.4 MODELLING CATEGORICAL OUTCOMES

The values of a categorical outcome (coded as 0 or 1) usually measure the presence or absence of particular event/characteristics and do not have a natural meaning in respect to location and scale. The distance between 0 and 1 in a categorical variable is not quantifiable in the same way as in a continuous variable where the measurements between points are of meaningful magnitude (Powers 2012).

Because of the uncertainty of the location and scale of the binary response, the focus is on modelling probability functions or transformations of probability functions. The transformations of the binary response variable are achieved by a variety of link functions (logit, probit, log-log, clog-log, etc). With the binary logistic regression for example what is predicted is not the dependent variable itself (which is the case with OLS via the identity link function) but is the logit of the dependent variable (the natural log of the odds that the dependent variable equals 1, conditional on the values of the predictors). The original values of the independent and dependent variables may be nonlinearly related while the link function of the dependent variable is always linearly related to the independent variables side of the equation (Garson 2013).

4.11.5 SINGLE LEVEL LOGIT

One approach to modelling a binary response Y(0,1) is the logit link function or the log-odds transformation. The logit function is often favoured because the estimated coefficients can be interpreted as odds ratios.

The following notations were taken from Rabe-Hesketh and Skrondal (2008, p.232) on generalized linear models, GLM (McCullagh and Nelder 1989). GLM is generic term for linear models for which the response variable is transformed in a nonlinear way (ex. logit link function).

When modelling binary response (as well as continuous) the interest is in the expectation (mean) of the response as a function of the predictors. The expectation for a binary response Y(0,1) is not the mean of Y but the probability that Y=1 given X predictors:

$$E(y_i|x_i) = Pr(y_i = 1|x_i)$$

When the response is continuous in nature the expectation of the response is modelled as a linear function of the predictors (here the regression lines increase or decrease without a limit as the predictors increase or decrease). For binary response this is problematic as the probability is not indefinite and must lie between 0 and 1. As a result a non-linear function is used in one of two ways:

$$Pr(y_i = 1|x_i) = h(\beta_1+\beta_2 x_i)$$

or

$$g\{Pr(y_i = 1|x_i)\} = \beta_1+\beta_2 x_i = v_i$$

Here $g(\cdot)$ is known as the *link function;* $h(\cdot)$ as the *inverse link function*, sometimes written as $g^{-1}(\cdot)$ and v_i is the *linear predictor*. The two formulas above are equivalent if $h(\cdot)$ is the inverse of $g(\cdot)$. The third element of the generalized linear model apart from the *link function* and the *linear predictor* is the *distribution* of the binary response given the predictors. The distribution for binary response (0,1) is specified as Bernoulli (π_i). The responses for different units (women) are assumed to be independent given the predictors.

If a logit link function is chosen, the above equations are expressed as:

$$Pr(y_i = 1|x_i) = logit^{-1}\ (\beta_1+\beta_2 x_i) \equiv \frac{\exp(\beta_1+\beta_2 x_i)}{1+\exp(\beta_1+\beta_2 x_i)} \qquad (1)$$

or

$$logit\{Pr(y_i = 1|x_i)\} \equiv \ln\ \{\frac{Pr(y_i = 1|x_i)}{1-Pr(y_i = 1|x_i)}\} = \beta_1+\beta_2 x_i \qquad (2)$$

The term $\{\frac{Pr(y_i = 1|x_i)}{1-Pr(y_i = 1|x_i)}\}$ is the Odds $(y_i = 1|x_i)$, i.e. the odds that $y_i = 1$ given x_i.

In other words what are the odds or the expected number of successes for each failure (the expected number of 1 responses for each 0 response). "*The natural log (ln) of the odds or the logit function of the probability is equated to the linear predictor. Correspondingly, (1) shows that the probability is given by the inverse logit function (sometimes called logistic function) of the linear predictor"* (Rabe-Hesketh and Skrondal 2008:232).

The relationship between odds and probabilities are given as:

$$\text{Odds} = \frac{Pr}{1-Pr} \text{ and } \text{Pr} = \frac{Odds}{1+Odds}$$

4.11.6 MULTILEVEL LOGIT MODEL – GENERALIZED LINEAR MIXED MODEL

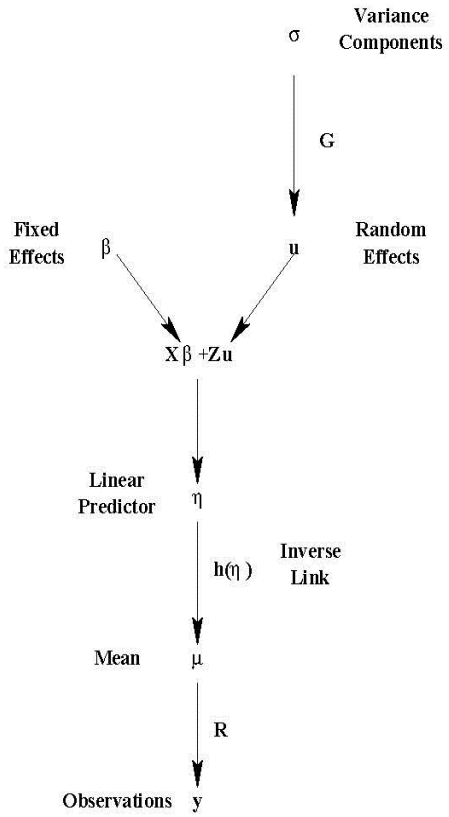
Multilevel modelling approach tests the multilevel theory. The multilevel theory hypothesizes that the variables at one level have an influence on the variables at another level (for example, maternity staffing (Level-2) may negatively affect the odds of giving birth by emergency CS (Level-1); i.e. by having more and highly skilled doctors and midwives, the odds of emergency CS are reduced).

A multilevel logit model belongs to the family of the Generalized Linear Mixed Models (GLMM), which incorporate nonlinear link function (logit, probit, etc.) of the dependent variable which is not continuous and not normally distributed. GLMM

combine linear mixed models which handle random effects and generalized linear models which handle nonnormal data via a link function and an exponential family distribution (binomial, Poisson, etc). *"GLMMs are the best tool for analyzing nonnormal data that involve random effects: all one has to do, in principle, is specify a distribution, link function and structure of the random effects*"(Bolker et al. 2009:127).

The graphic presentation (Figure 6) of a GLMM is taken from Stephen D. Kachman (2000:61) http://statistics.unl.edu/faculty/steve/glmm/paper.pdf

Figure 6: Graphic presentation of GLMM, adopted from Kachman 2000:61



In Figure 6,

- X and Z are known as design matrices;

- β are fixed effect (e.g. age, ethnicity, IMD, etc.);
- u are the random effects u~N(0,G);
- R and G are covariance matrices which depend on a set of unknown variance components;
- y is a vector of observations (i.e. emergency CS yes/no) for which the conditional distribution given the random effects has a mean μ (e.g. mean emergency CS rate) and a covariance matrix R (e.g. variance of emergency CS status is μ(1-μ)); the mean depends on the linear predictor via the inverse link function; while the covariance matrix R depends on the mean via a variance function
- η is the linear predictor; and
- h(η) is the inverse link function.

The basic GLMM for binary response with a random intercept only (or a 2-Level multilevel logit model for binary response), in terms of expectation for the binary response Y(0,1), is expressed as:

$$E(y_{ij}) = Pr(y_{ij} = 1) = logit^{-1}(\beta_0 + \beta_1 x_{ij} + u_j)$$
$$\equiv \frac{\exp(\beta_0 + \beta_1 x_{ij} + u_j)}{1 + \exp(\beta_0 + \beta_1 x_{ij} + u_j)}$$

or

$$logit\{Pr(y_{ij} = 1)\} \equiv \ln\{\frac{Pr(y_{ij}=1)}{1-Pr(y_{ij}=1)}\} = \beta_0 + \beta_1 x_{ij} + u_j \quad (3a)$$

where y_{ij} takes the value 0 or 1 for each individual i in group j (for example: 1= emergency CS; 0 = not emergency CS; for each woman i in trust j); Pr (y_{ij}) is the probability of observing the response $y_{ij} = 1$ for woman i in trust j; β_0 is the constant term 'intercept'; β_1 is the coefficient of the individual level predictor x_{ij}

(which may be continuous, categorical or dummy variables) and u_j is the Level- 2 residual.

The fixed part of the model in (3a) is interpreted as follows:

- β_0 is the overall intercept representing the baseline log odds (or logit) that y_{ij}= 1 when $x_{ij} = 0$ and $u_j = 0$;
- β_1 quantifies the change in the logit corresponding to a unit change in X for individuals in the same group (same value of u_j ,i.e. for women in the same trust). β_1 is also interpreted as the cluster-specific effect of X;
- exp (β_1) is the odds ratio, comparing odds for women spaced one unit apart on X but in the same trust.

The random part of the model in (3a) is interpreted as follows:

- u_j is interpreted as the effect of being in trust j on the log-odds that y_{ij}= 1. It is the group random effect or Level-2 residual;
- $\beta_0 + u_j$ is the intercept for a given group j and will be higher or lower than the overall intercept depending on u_j being greater or smaller than 0;
- estimates and confidence intervals for u_j can be calculated;
- u_j is assumed to be normally distributed with 0 mean and variance σ_u^2;
- σ_u^2 is the Level-2 (residual) variance, or the between-trust variance in the log-odds that y_{ij}= 1 after accounting for x_{ij}. It is the unexplained Level-2 variance in presence of predictors;
- because the part of the model shown by equation (3a) is for a binary response variable, there is no Level-1 residual because its variance in a Bernoulli distribution is a function of the population proportion (mean) and therefore cannot be estimated separately (Heck et al. 2012). A scale factor of 1.0 for the variance is used in most software programmes to establish a metric for the

linear predictor (Heck et al. 2012). The variance of a logistic distribution with a scale factor of 1.0 is $\pi^2/3$~3.29 and is therefore assumed fixed under the log-odds transformation (Snijders and Bosker 1999; Raudenbush and Bryk 2002). The implication of fixing the residual variance in such a way for multilevel binary response models is that the Level-1 variance is not estimated in these models. The residual u_j and its variance σ_u^2 are specified and estimated at Level-2 instead. A consequence of that is that, when a more complex model is built from an intercept only model (Null), it is difficult to ascertain how much variance is accounted for by successive models with more predictors, because the underlying outcome is rescaled to 1.0 each time. This rescaling also affects the Level-2 variance and thus making it difficult to assess (and therefore interpret) changes in variance accounted for between successive models as will be the case with multiple regression models (Heck 2012). On a positive side the fixing of the Level-1 variance facilitates the estimation of the Variance Partition Coefficient (VPC) in a multilevel binary response model, i.e. the proportion of the variance in the binary outcome that lies between groups (trusts) or:

$$VPC = \frac{\sigma_u^2}{3.29 + \sigma_u^2}$$

The log-odds are not easily interpreted. The log-odds coefficients though can be transformed into odds ratios (which are not the same as odds of success) by e^{β}, where e=2.71828. The odds ratios are interpreted as the change in the odds of the outcome due to a unit change in a predictor X, while holding all else constant, i.e. they tell us about the strength of the association between a predictor and an outcome.

The basic model could be modified to include more levels by including error terms at all levels above the individuals. The model shown above is of the form of a 'random intercept model'. It could be extended to include more predictors at Level-1 and Level-2, cross-level interactions and random slopes (coefficients) for the predictors, and hence take the form of a 'random slopes' model.

4.11.7 LATENT PRESENTATION MODEL AND VARIANCE PARTITION COEFFICIENT (VPC)

Some disciplines (especially economics and psychology) prefer to present binary response models in terms of a latent continuous variable y*. In clinical research, outcomes that cannot be measured directly such as mental health, physical disability or quality of life, are known as latent variable. The assumption is that there is a latent continuous variable y* that underlines the observed binary y, representing for example the propensity for emergency CS (though the outcome is clear, i.e. a woman either had an emergency CS or not, it was not possible to observe directly how the decision to perform emergency CS was taken, at what stage of labour and was it because of specific maternal/baby risk conditions that developed during labour, deficiencies in care - inexperienced staff or shortages of staff, institutional practices, etc. such that:

$$y_{ij} = \begin{cases} 1 \; if \; y_{ij}^{*} \geq 0 \\ 0 \; if \; y_{ij}^{*} < 0 \end{cases}$$

The threshold model for the latent response y_{ij}^{*} is specified as:

$$y_{ij}^{*} = \beta_0 + \beta_1 x_{ij} + u_j + e_{ij}^{*} \text{ , where}$$

$e_{ij}^{*} \sim$ N(0; 1) in a probit model or

$e_{ij}^{*} \sim$ N(0; $\pi^2/3$) in a logit model, i.e. e_{ij}^{*} follows standard logistic distribution given x_{ij} (with mean 0 and variance $\pi^2/3 \sim 3.29$) in a logit model, meaning that the Level-1 variance e_{ij}^{*} is fixed.

Latent presentation model allows for the calculation of the Variance Partition Coefficient (VPC) in the following way:

$$VPC = \frac{\sigma_u^2}{\sigma_{e*}^2 + \sigma_u^2}$$

Where σ_{e*}^2 = 3.29 for the logit model.

Variance partition coefficient (VPC) measures the proportion of the total residual variance due to the differences between groups (trusts). For example a VPC of 0.2 (is the correlation between two randomly selected individuals from the same group) and could also be interpreted as 20% of the variation in the outcome variable is between groups and 80% is within groups. VPC and ICC are the same in simple multilevel models for continuous outcomes. In the case of multilevel models with a binary outcome the VPC is interpreted as "*the proportion of the total residual variance in the propensity to be in response category 1 that is due to differences between groups"* (C 7.2: Latent Variable Representation of a Random Intercept Model for Binary Responses, CMM Bristol online tutorials).

4.11.8 ESTIMATION PROCEDURES

The following is information taken from: IBM SPSS Statistics (2010) *Make Smarter Decisions with Your Nested Data. Using GLMM with continuous and categorical targets. IBM Software Business Analitics (2010)*.

Because of the two probability distributions[103] in GLMM, it is difficult to estimate the parameters (these include fixed effects parameters, variances and covariances in G – the random effect covariance matrix). There are three estimation methods suggested in the statistical literature: linearization-based method; integral approximation method and Bayesian method.

[103] One is for the conditional target variable, the other is for the random effects.

1. Linearization based methods – these approximate the GLMM, via Taylor series expansion and transforms it into linear mixed pseudo model (LMM). Reliable estimation methods for LMMs exist. These methods are called different names in the statistical literature depending on the motivation, i.e. pseudo likelihood approach (PL); penalized quasi-likelihood (PQL), marginal quasi-likelihood (MQL) and so on.
2. Integral approximation methods – these use iterative techniques to approximate the log likelihood and numerically optimize the approximated function. Obtaining the log likelihood values via numerical integration techniques, such as Gaussian-Hermite quadrature or Laplace approximation over the random effects distribution is difficult.
3. Bayesian methods – these use the Markov chain Monte Carlo (MCMC) for GLMMs fitting. These are computationally intensive as they use prior distributions of parameters to obtain posterior distributions of parameters and then estimate the parameters.

IBM SPSS 22 uses linearization-based estimation method (PQL). The advantages of this method are that it's easier to implement; run faster; and can be used in more complex model structures (for large number of random effects or correlated errors). The shortcomings are that this method produces asymptotically biased estimates, especially for binary targets. IBM SPSS Statistics (2010) suggests that GLMM should be used to understand the relationship between target and predictors but not for prediction; and the focus should be on the interpretation of the fixed effects, not on the variances and covariances. However it is known that the random effects have an effect on the significance and interpretation of the fixed effects. Model diagnostics in GLMMs are also a topic of continuing research (IBM SPSS Statistics 2010).

Different software programs use different iterative methods to estimate the parameters in the multilevel models. Multilevel models for continuous response are usually estimated with a maximum likelihood (ML) procedure. For binary responses there is a choice of options. A direct ML via numerical quadrature is available in SAS and STATA SE 12 but not in MLwiN 2.26 or IBM SPSS 22. MLwiN 2.26 uses

quasi-likelihood procedures (Marginal Quasi Likelihood (MQL) and Penalized Quasi Likelihood (PQL)) and Markov chain Monte Carlo methods (MCMC).

A paper by Rodriguez and Goldman (2001) which reviewed several types of approximate procedures, revealed that these methods could produce different results. The recommendations from the Centre for Multilevel Modelling (CMM online tutorials, Bristol) is that quasi-likelihood methods (MQL and PQL) are quick and useful for screening but could produce biased results if the cluster sizes are small; ML methods are recommended for simple models but could be slow to reach convergence if some of the random effects are correlated; and MCMC are flexible and becoming more computationally feasible thus recommended for use in MLwiN.

Penalized Quasi Likelihood (PQL) is the approximation procedures used through the multilevel analyses in this study.

4.11.9 OVERDISPERSION AND UNDERDISPERSION

As described earlier Level-1 variance in a multilevel logit model cannot be estimated separately and is fixed at a scale factor of 1.0 in most software programmes. Hox (2010) warns that the assumption behind this scale factor of 1.0 is that Level-1 errors follow the binomial error distribution exactly (Heck 2012:94). In data with cases of extreme outliers, extremely small groups and if a whole level is omitted, this exact fixing may lead to overdispersion (when the observed variance is larger than would be expected) or underdispersion (when observed variance is smaller than would be expected) (Heck 2012).

Wald Z and χ^2 test the null hypothesis of no effect (the parameters or combination of parameters are scaled by their standard errors; the resulting test statistics is compared to 0). These two tests are only appropriate for GLMMs without overdispersion. In case of overdispersion two other tests are preferred - Wald t and F. These account for the uncertainty in the estimates because of overdispersion (Bolker et al. 2009).

The uncertainty in the estimates is determined by the number of residual degrees of freedom. These are difficult to calculate. The number of the parameters used by a random effect lies between 1 (i.e. a single standard deviation parameter) and N-1 (i.e. one parameter for each additional level of the random effect). For random effects, Wald Z and χ^2 suffer from boundary effects because the null values of the parameters '*lie at the edge of their feasible range'* meaning that the '*standard deviations can only be greater and not less than their null hypothesis value of zero'* (Bolker et al. 2009).

IBM SPSS uses Wald Z statistics to test whether the between trust variance is statistically different from 0; and F tests for fixed effects. It also uses Satterthwaite approximation to adjust for denominator degrees of freedom when calculating the fixed effects. Satterthwaite approximation was recommended in GLMM IBM SPSS 22 for smaller sample size, unbalanced data or complicated covariance structure such as unstructured.

4.12 THE MODELS

Multilevel logistic regression analyses were applied to the three binary indicators: emergency caesarean section, normal birth and instrumental delivery. This approach accounts for the clustering of the data and the within-trust correlation of the observed outcomes; estimates the between-trust variation in outcomes; and examines the hypothesis that trusts and workforce differences in the provision of maternity care contribute to the variations in the outcomes between trusts. The percentages of the total variation in outcomes attributable to between trusts variation were calculated.

The predictors at trust level were – number of FTE medical and non-medical staff per birth (consultant O&G FTE per birth; doctor FTE per birth; midwife FTE per birth; HCA FTE per birth); trust characteristics (London trust, teaching trust, foundation trust); trust configuration (OU/AMU/FMU).

Individual level predictors included socio-demographic patient characteristics (maternal age; index of multiple deprivation (IMD) in quintiles; ethnicity); NICE (2007) clinical risk; birth weight and gestational age.

4.12.1 MODELLING STAGES

Before considering the inclusion of predictors, it is of interest to know the levels at which significant variation exists. With a two-level model, it is generally assumed that within-trust variation is present. That is not necessarily the case with between trust variations. So it is important to start with examining the intercept variability and later to consider between-group slope variation as well (Bliese 2013).

For all the models only a random intercept was modelled. The different modelling stages included:

- No Level-1 predictors – This is the Null model, also called the unconditional model. The Null model models the outcome without predictors apart from the random effect of the Level-2 grouping variable (trust).
- Level-1 predictors only – The outcome is predicted from the fixed effects of Level-1 predictors and a random effect of Level-2 grouping variable (trust).
- Both Level-1 and Level-2 predictors - The outcome is predicted from the fixed effects of Level-1 and Level-2 predictors and a random effect of the Level-2 grouping variable (trust).

Multilevel logistic models were fitted separately for each outcome. The modelling stages were determined by the main research questions. Following is an explanation for the modelling steps taken to answer each of the four research questions.

Q1: What is the extent of between trusts variations in emergency caesarean section, instrumental delivery and normal birth?

A multilevel modelling approach was needed to answer this question. More specifically, Null models were fitted, e.g. two-level random intercept models with no

predictors to explore the extent (if any) of between-trusts variations separately in emergency CS, instrumental delivery and normal birth.

Q2: Do women's socio-demographic characteristics and/or clinical risk affect their probability of emergency CS, instrumental delivery and normal birth?

A single level analysis could answer this question, assuming that the women were randomly sampled and that higher level clusters of women did not exist. In this study the two-level random intercept models (the Null models) were extended by adding individual-level covariates: women's age; ethnicity; IMD in quintiles; NICE (2007) risk, gestational age; and baby's birth weight to examine the strength of the relationships between Level-1 covariates and the three outcomes.

Q3: Do NHS trusts configuration and staffing levels influence the probabilities of emergency CS, instrumental delivery and normal birth and do they explain any of the variations in outcomes between trusts?

To answer this question and to explore the contextual effects on the outcomes, Level-2 (trust level) predictors were added to the random intercept and individual-level covariates models in two blocks. The first block included, London trust; Foundation trust; University Hospital; OU/AMU/FMU and the second the standardized staffing ratios of consultants O&G; doctors, midwives and healthcare assistants (HCA), all FTE per birth. An additional aim was to find out if the Level-2 predictors explained any of the Level-2 variations in the three outcomes between trusts.

Q4: Are the effects (if any) of staffing levels on the outcomes different for low-risk and high-risk women?

Cross-level interaction effects between each staffing group and individual-level risk were explored. If the interaction is significant, it will be considered to have a moderation effect, e.g. significant coefficients will indicate that the effects of staffing on the outcomes vary with women's clinical risk.

4.12.2 SOFTWARE

IBM SPSS Statistics 22 was used for the preparation of the data (matching, recoding, initial exploration and descriptive statistics) and the GLMM in IBM SPSS 22 was used for the multilevel modelling. Some of the results were also checked using software packages MLwiN 2.26 (Centre for Multilevel Modelling, University of Bristol) and STATA SE 12 (StataCorp LP).

5 CHAPTER 5 RESULTS

This Chapter presents the modelling results.

The study population comprised women aged 15-44, who were nulliparous and had a term, singleton, live birth (n=261,481 for emergency caesarean section and instrumental deliveries in 143 NHS trusts; and n=214,949 for normal birth in 129 NHS trusts). Individual covariates included maternal age, ethnicity, level of deprivation (IMD), gestational age, birth weight and a composite measure of clinical risk based on the NICE 2007 intrapartum guideline definition. The trust level predictors were FTE/birth ratios for consultant obstetricians and gynaecologists, doctors (as defined in Chapter 4), midwives, healthcare assistants, London/Teaching/Foundation trust and trust configurations related to any combination of obstetric/alongside-midwifery/free-standing midwifery units within trusts.

Please note that hierarchical and multilevel terms were used synonymously (the first is mainly used in educational sciences and the second is used in social sciences)

5.1 RESULTS

5.1.1 DESCRIPTIVE STATISTICS HES 2010/11

5.1.1.1 HES POPULATION DESCRIPTIVES

From the 657,480 deliveries in the 143 trusts, 14.7% were emergency CS; 33.6% were normal and 12.4% were instrumental deliveries (which included use of forceps and ventouse). A higher proportion of women were classified as NICE high risk (52.1%) than NICE low risk (47.9%); 42.9 % were nulliparous; 98.3% had a singleton birth; 88.4% had a live birth; 74.4% were aged 20-35; 74.4% were white; 73.9% had an infant with gestational age between 37 and 41 weeks and 82.5% had infants weighing between 2500 and 4500 grams (Table 5.1-1). Gestational age was the variable with most missing data (16.8%), followed by live birth (11.2%) and

infant birth weight (10.2%). Normal birth outcome had 16.5% missing data because of its composite nature (mainly due to the use of regional anaesthetic, episiotomy and induction variables, which had respectively 30.8%, 24.9% and 11.1% missing data). The overall missing data in the HES population for emergency CS and instrumental delivery was 0.1% (840 cases) and for normal birth was 16.5% (108,740).

Table 5.1-1: Descriptive Statistics; HES 143 trusts; 657,480 cases; April 2010/March 2011

Variables	Categories	Count	Column %
Emergency CS	Unknown	840	.1
	No	559754	85.1
	Yes	96886	14.7
Instrumental Delivery	Unknown	840	.1
	No	574810	87.4
	Yes	81830	12.4
Normal Birth	Unknown	108740	16.5
	No	327741	49.8
	Yes	220999	33.6
Parity	Multiparous	375496	57.1
	Nulliparous	281984	42.9
Multiple Births	Unknown	368	.1
	Singleton	646279	98.3
	Multiple	10833	1.6
Live Births	Unknown	73785	11.2
	Live Birth	581187	88.4
	Still Birth	2508	.4
NICE Any Risk	Lower Risk	314838	47.9
	Higher Risk	342642	52.1
Maternal Age (years)	15-19	35389	5.4
	20-24	124157	18.9
	25-29	180432	27.4
	30-34	184985	28.1
	35-39	104997	16.0
	40-44	24075	3.7
	Unknown	1868	.3
	Other Ages	1577	.2

Ethnicity	White	489098	74.4
	Asian	69398	10.6
	Afro-Caribbean	36150	5.5
	Mixed, Chinese, Other	33237	5.1
	Unknown	29597	4.5
Level of Deprivation (IMD)	1 Least Deprived	130396	19.8
	2	130409	19.8
	3	130393	19.8
	4	130411	19.8
	5 Most Deprived	130399	19.8
	Unknown	5472	.8
Gestational Age (weeks)	37w<=GA<=41w	485726	73.9
	GA>41w	24254	3.7
	GA<37w	37089	5.6
	Unknown	110411	16.8
Birth Weight (g)	2500g<=BW<=4500g	542677	82.5
	BW<2500g	37807	5.8
	BW>4500g	9992	1.5
	Unknown	67004	10.2

Source: HES, April 2010/March 2011

Table 5.1-2 presents the row percentages of demographic and clinical risk factors by the three birth outcomes for all the deliveries in HES[104]. Older women were more likely to experience emergency CS (19.5% for the 40-44 years old vs. 11.3% of the 15-19 years old), while the 15-19 years (13.4%) and 30-34 years (13.2%) age groups were most likely to have instrumental deliveries. The 20-24 years age group was most likely to have a normal birth. White women were least likely to have an emergency CS (13.9%) and most likely to have an instrumental delivery (12.8%); while Afro-Caribbean women were most likely to have an emergency CS (20.7%) and normal birth (42.4%) and least likely to deliver with an instrument (6.4%). Women with the lowest IMD had the lowest rate of instrumental delivery (9.8%) and

[104] Emergency CS and Instrumental Delivery were based on 656,640 cases (840 missing cases for emergency CS and instrumental delivery were excluded), while Normal Birth was based on 548,740 cases (108,740 missing cases for normal birth were excluded as normal birth is a composite measure).

highest rate of normal birth (44.9%). More high risk women delivered by emergency CS (20.3%); while a higher percentage of the low risk women had instrumental (13.6%) and normal (58.4%) birth. 30.7% of women with low infant birth weight births (<2500g) delivered by emergency CS; while the highest percentages of instrumental (12.8%) and normal births (41.3%) were among women with normal infant birth weight (2500-4500g). 42.1% of women who were between 37-41 weeks gestation at the time of birth had a normal birth and of those with pre-term babies (<37w), 31% had an emergency CS.

Table 5.1-2: Demographic and clinical risk factors by birth outcomes HES April 2010/March 2011

		Emergency CS		Instrumental Del		Normal Birth	
		656,640 cases 143 trusts		656,640 cases 143 trusts		548,740 cases 129 trusts	
		Count	Row %	Count	Row %	Count	Row %
Maternal Age (years)	15-19	4004	11.3	4733	13.4	13385	44.1
	20-24	15468	12.5	14827	12.0	47585	45.1
	25-29	25316	14.0	23010	12.8	63998	42.1
	30-34	28838	15.6	24453	13.2	58863	38.5
	35-39	17840	17.0	12088	11.5	30469	35.5
	40-44	4698	19.5	2453	10.2	5844	30.0
	Unknown	370	20.2	125	6.8	565	37.1
	Other Ages	352	22.3	141	9.0	290	23.2
Ethnicity	White	68080	13.9	62312	12.8	164307	39.8
	Asian	11356	16.4	8217	11.8	24008	42.3
	Afro-Caribbean	7459	20.7	2311	6.4	12099	42.4
	Mixed, Chinese, Other	5285	15.9	4134	12.5	11008	42.0
	Unknown	4706	15.9	4856	16.4	9577	39.4
Level of Deprivation (IMD)	1 Least Deprived	19117	14.7	18513	14.2	40031	36.9
	2	19316	14.8	17824	13.7	41018	38.0
	3	19394	14.9	17009	13.1	42449	39.6
	4	19652	15.1	15105	11.6	45585	42.0
	5 Most Deprived	18416	14.1	12758	9.8	50421	44.9
	Unknown	991	18.1	621	11.4	1495	36.6
NICE Risk	Lower Risk	27433	8.7	42652	13.6	154481	58.4

	Higher Risk	69453	20.3	39178	11.4	66518	23.4
Gestational Age (weeks)	37w<=GA<=41w	62790	12.9	60633	12.5	183883	42.1
	GA>41w	6237	25.7	4748	19.6	3793	17.4
	GA<37w	11482	31.0	2556	6.9	10271	31.6
	Unknown	16377	14.9	13893	12.6	23052	40.2
Birth Weight (g)	2500g<=BW<=4500g	72747	13.4	69609	12.8	202680	41.3
	BW<2500g	11606	30.7	2705	7.2	9765	29.2
	BW>4500g	2570	25.7	1105	11.1	2486	27.5
	Unknown	9963	15.0	8411	12.6	6068	37.7

Source: HES, April 2010/March 2011

5.1.1.2 HES SAMPLE SELECTION DESCRIPTIVES

The sample selection section presents the descriptive statistics for women aged 15-44, who were nulliparous and had a term (>=37 weeks), singleton, live birth. The overall sample comprised of 261,895 deliveries in 143 trusts, among such women.

This sample differed from the HES population in the following ways:

- a higher proportion of women had an emergency CS (19.2% vs 14.7%); a higher proportion also had instrumental delivery (21.8% vs 12.4%) and a smaller proportion had normal birth (24.7 vs 33.6%);
- the sample had a higher proportion (10.2%) of women in the youngest age group (15-19) compared to 5.4% in the HES population; and a lower proportion of women aged >35 (13% vs 19.7%);
- it had a smaller proportion of women living in a deprived area based on IMD (16.9% vs. 20.2% in the least deprived quintile);
- a higher proportion of women were classified as low risk than high risk based on NICE criteria (57.3% vs. 42.7%). In comparison, a higher proportion of women in HES population were classified as NICE high risk than low risk (52.1% vs. 47.9%);

- the sample had marginally higher proportion of women with GA>41w and a lower proportion of women who had babies weighing less than 2500g;
- the sample also had marginally lower proportion of Asian and Afro-Caribbean women.

Table 5.1-3 presents the detailed descriptive statistics for the sample of 15-44 years old, nulliparous, at term with singleton, live births.

Table 5.1-3: Descriptive Statistics, HES sample of 261,895 women; aged 15-44 years; nulliparous; at term (>=37w) with singleton, live births in 143 trusts

Variables	Categories	Count	Column %
Emergency CS	Unknown	427	.2
	No	210878	80.5
	Yes	50590	19.3
Instrumental Delivery	Unknown	427	.2
	No	204250	78.0
	Yes	57218	21.8
Normal Birth	Unknown	46957	17.9
	No	150193	57.3
	Yes	64745	24.7
Nulliparous	Yes	261895	100.0
Singleton	Unknown	167	.1
	Yes	261728	99.9
Live Births	Unknown	34801	13.3
	Yes	227094	86.7
NICE Any Risk	Lower Risk	150072	57.3
	Higher Risk	111823	42.7
Maternal Age (years)	15-19	26682	10.2
	20-24	61149	23.3
	25-29	73666	28.1
	30-34	65539	25.0
	35-39	28088	10.7
	40-44	5938	2.3
	Unknown	833	.3
Ethnicity	White	195191	74.5
	Asian	22623	8.6

	Afro-Caribbean	11362	4.3
	Mixed Chinese Other	14020	5.4
	Unknown	18699	7.1
Level of Deprivation (IMD)	1 Least Deprived	52926	20.2
	2	54808	20.9
	3	55316	21.1
	4	51912	19.8
	5 Most Deprived	44357	16.9
	Unknown	2576	1.0
Gestational Age (weeks)	37w<=GA<=41w	199652	76.2
	GA>41w	13951	5.3
	Unknown	48292	18.4
Birth Weight (g)	2500g<=BW<=4500g	219157	83.7
	BW<2500g	6996	2.7
	BW>4500g	3114	1.2
	Unknown	32628	12.5

Source: HES, April 2010/March 2011

The sample was further divided into two sub-samples, one which included only the valid cases for emergency CS and instrumental deliveries (427 cases or 0.2% were excluded) and a second one which contained only the valid normal birth cases (46,957 cases or 17.9% were excluded). The final sub-sample for valid emergency CS and instrumental deliveries comprised of 261,468 deliveries in the 143 trusts; and the sub-sample for valid normal deliveries was for 214,938 deliveries in 129 trusts; all deliveries from women who were aged 15-44 years, nulliparous, with singleton, live births at term (>=37 weeks). Table 5.1-4 shows the row percentages of demographic and clinical risk factors by the three birth outcomes (valid cases only) for the two sub-samples of women.

Table 5.1-4: Demographic and clinical risk factors by birth outcomes for women aged 15-44 years, nulliparous, with singleton, live births at term (>=37 weeks)

		Emergency CS		Instrumental Delivery		Normal Birth	
		261,468 cases 143 trusts		261,468 cases 143 trusts		214,938 cases 129 trusts	
		Count	Row %	Count	Row %	Count	Row %
Maternal Age (years)	15-19	3016	11.3	4077	15.3	9473	41.5
	20-24	9429	15.4	11511	18.8	18483	35.8
	25-29	14165	19.3	16905	23.0	18172	29.7
	30-34	14603	22.3	16603	25.4	12962	24.8
	35-39	7487	26.7	6818	24.3	4650	21.2
	40-44	1722	29.0	1210	20.4	807	17.9
	Unknown	168	21.0	94	11.8	198	32.8
Ethnicity	White	36086	18.5	43657	22.4	48918	30.0
	Asian	5206	23.0	5251	23.2	4648	25.9
	Afro-Caribbean	2998	26.4	1339	11.8	2842	33.6
	Mixed Chinese Other	2881	20.6	2911	20.8	3284	31.2
	Unknown	3419	18.3	4060	21.8	5053	33.4
IMD	1 Least Deprived	10340	19.6	12864	24.3	11844	27.3
	2	10759	19.7	12657	23.1	12818	28.7
	3	10713	19.4	12248	22.2	13324	29.9
	4	10036	19.4	10560	20.4	13451	31.8
	5 Most Deprived	8220	18.6	8421	19.0	12755	33.3
	Unknown	522	20.3	468	18.2	553	30.8
NICE Any Risk	Lower Risk	19705	13.2	33236	22.2	52811	42.4
	Higher Risk	30885	27.6	23982	21.5	11934	13.2
Gestational Age	37w<=GA<=41w	36500	18.3	43317	21.7	56740	31.5
	GA>41w	4657	33.4	3676	26.4	1272	10.1
	Unknown	9433	19.6	10225	21.3	6733	30.5
Birth Weight	2500g<=BW<=4500g	41019	18.7	48714	22.2	60402	30.5
	BW<2500g	1815	26.0	1065	15.2	1712	27.5
	BW>4500g	1430	45.9	656	21.1	291	10.4
	Unknown	6326	19.5	6783	20.9	2340	29.4

Source: HES, April 2010/March 2011

The row percentages of demographic and clinical risk factors by birth outcomes in Table 5.1-4 (sub-sample) were compared to the row percentages in Table 5.1-2 (HES population). In the sub-sample of nulliparous women, more women in the high risk (HR) group experienced emergency CS (27.6%), compared to 20.3% in the overall HES population. A higher proportion of low risk (LR) women in the sub-sample had an instrumental delivery (22.2%) compared to the overall rate for LR of 13.6% in the HES population; and a smaller proportion of LR women in the sub-sample had a normal birth (42.4%) compared to the overall rate for LR of 58.4%. The overall emergency CS and instrumental delivery rates in HES were 14.7% and 12.4% respectively (Table 5.1-1). The sub-sample was similar to the overall population in terms of maternal age and deprivation patterns across all outcomes.

In terms of ethnicity, Asian women had the highest proportion of instrumental deliveries (23.2%) in the sub-sample, while white women had the highest rate (12.8%) in the HES population. Higher percentages of women in the sub-sample, with a pregnancy gestation greater than 42 weeks, experienced an emergency CS (33.4%) and instrumental deliveries (26.4%) compared to the same gestation group in the population (25.7% and 19.6% respectively), while a smaller proportion of women with pregnancy gestation of 37-41 weeks in the sub-sample had a normal birth (31.5%) compared to the same group in the HES population (42.1%). More women delivering babies with a normal birth weight (2500-4500g) had an emergency CS (18.7%) or instrumental delivery (21.7%) in the sub-sample compared to the HES population (13.4% and 12.8% respectively).

5.1.1.3 MODE OF BIRTH ACCROSS TRUSTS

In the two sub-samples of women aged 15-44, nulliparous, with singleton, live births at term (>=37 weeks):

- Rates of Emergency CS for the 143 trusts ranged from 11.3% to 31.1%;
- Rates of instrumental delivery ranged from 11.4% to 31.7%;

- Rates of normal birth for 127 trusts ranged from 20.4% to 43.5% and for the remaining two trusts the values were 100% (based on one case) and 58.8% (17 cases only). These two trusts were additionally dropped at the multilevel modelling stage of normal birth due to the very small cell count when modelling normal birth outcome.

5.1.2 STAFFING FTE/BIRTH RATIOS

The raw numbers on FTE and HC for each maternity medical and non-medical staff group were presented earlier in Table 4.5-2 and Table 4.5-3. Table 5.1-5 presents the descriptive statistics of maternity staff FTE/birth ratios (consultants, doctors, midwives and HCAs in the 143 trusts). The FTE per birth ratio is easier to understand if multiplied by 1000 and interpreted as FTE in each staff group per 1000 births (i.e. the median trust had 30.6 FTE midwives per 1000 births - one trust had 11.1 FTE midwives per 1000 birth, while another had 45.6 FTE midwives per 1000 births).

Table 5.1-5: NHS IC Maternity Workforce Data, at trust level, September 2010

		Consultants FTE/birth	Registrars FTE/birth	Junior Doctors FTE/birth	Doctors FTE/birth	Midwives FTE/birth	HCA FTE/birth
N	Valid	143	143	137	143	143	143
	Missing	0	0	6	0	0	0
Mean		.0026	.0045	.0011	.0056	.0304	.0098
Median		.0026	.0044	.0010	.0054	.0306	.0095
Std. Deviation		.0006	.0015	.0010	.0017	.0050	.0036
Skewness		.3081	1.4302	2.7004	1.0424	.2132	1.5946
Std. Error of Skewness		.2027	.2027	.2070	.2027	.2027	.2027
Kurtosis		.2835	6.1452	11.4662	3.6975	2.0588	6.5719
Std. Error of Kurtosis		.4028	.4028	.4112	.4028	.4028	.4028
Minimum		.0007	.0005	.0001	.0005	.0111	.0010
Maximum		.0046	.0128	.0070	.0138	.0456	.0288
Percentiles	25	.0021	.0036	.0005	.0045	.0272	.0079
	50	.0026	.0044	.0010	.0054	.0306	.0095
	75	.0031	.0051	.0014	.0064	.0331	.0112
IQR (Q3-Q1)		.0010	.0015	.0009	.0019	.0059	.0033

Source: NHS IC Maternity Workforce Census Data, September 2010

The skewness estimates divided by their standard errors (SE) (Table 5.1-5), suggested that the distributions of the ratios of Consultant FTE/birth and Midwives FTE/birth were approximately symmetric (skewness/SE<2 in both cases). The skewness/SE estimates of registrars FTE/birth, junior doctors FTE/birth, and HCAs FTE/birth ratios were all greater than 2, and therefore highly and positively skewed. The kurtosis estimate divided by its SE was less than 2 for consultants FTE/birth ratio, showing platykurtic distribution, which means lower and broader central peaks and shorter and thinner tails (compared to normal distribution). The kurtosis/SE estimates for the rest of the staffing/birth ratios were all greater than 2, meaning excess kurtosis or leptokurtic distributions. Compared to normal distribution the plots show higher and narrower peaks and fatter and longer tails. Balanda and MacGillivray (1988) in a critical review of kurtosis explain that increasing kurtosis is

"associated with the movement of probability mass from the shoulders of a distribution into its center and tails" (cited in Brown 2012). An online tutorial by Brown (2012) says the same in different words "*higher kurtosis means more of the variability is due to a few extreme differences from the mean, rather than a lot of modest differences from the mean*" (p.8). This is helpful to know when investigating how trust differences in staffing/birth ratios may explain the variations in the three outcomes across trusts. It seemed that for a critical mass of trusts there was little variability in the doctors, midwives and HCA FTE/birth ratios and that few trusts had extreme differences from the overall mean values of doctors, midwives and HCA FTE/birth ratios.

Standardized staffing FTE/birth ratios were used in the multilevel models. The standardization was done at the level of the trust (Table 5.1-6) before the staffing variables were matched to the HES individual level dataset.

Table 5.1-6: Standardized staff FTE/birth ratios at trust level

	N	Minimum	Maximum	Mean	Std. Deviation
ZConsultants FTE/birth	143	-3.004	3.200	0	1
ZRegistrars FTE/birth	143	-2.645	5.469	0	1
ZJunior Doctors FTE/birth	137	-1.045	6.111	0	1
ZDoctors FTE/birth	143	-2.978	4.826	0	1
ZMidwives FTE/birth	143	-3.897	3.074	0	1
ZHCA FTE/birth	143	-2.440	5.290	0	1

Kreft et al. (1995) provide a good discussion of the consequences of various centring schemes and justification for their use. Mean centring was recommended in multilevel models for stabilizing the data.

Workforce, HES and trusts characteristics datasets were matched at trust level using unique trusts identifiers. Hence the matched data allowed us to fit a 2-Level multilevel logistic regression model – Level-1 (mothers) nested in Level-2 (trusts) to model different birth outcomes.

5.1.3 EMERGENCY CAESAREAN SECTION – MULTILEVEL MODEL

As described earlier, the modelling stages were determined by the research questions.

Q1: What is the extent (if any) of between-trusts variation in emergency caesarean section?

To answer this question, a Null model was fitted first, e.g. a two-level random intercept model with no predictors to explore the extent (if any) of between-trusts variations in emergency CS.

For the Null model "t*he intercept β_{0j} consists of two components: a fixed effect β_0 shared by all trusts, and a random effect u_{0j} specific to trust j. The random effect is assumed to follow a normal distribution with covariance matrix Ω_u, which in this model contains just one element, the between-trust variance σ^2_{u0}"* (Fiona Steele, Centre for Multilevel Modelling, Bristol: Module 7 (Practical): Multilevel Models for Binary response).

All the models are run by using quasi-likelihood procedure. IBM SPSS provides only PQL penalized or predictive quasi-likelihood (PQL) estimates.

The PQL estimates of the Null model (Table 5.1-7) for emergency CS showed that the log-odds of experiencing emergency CS in an average trust (one with u_{0j}=0) was estimated at: -1.441 (SE 0.018). The intercept for trust j was: -1.441+ u_{0j} and the variance σ^2_{u0} of u_{0j} was estimated at: 0.040 (SE 0.005). SPSS used Wald Z statistics to test whether the variance was statistically different from 0. Wald Z statistics tests the null hypothesis that $\sigma^2_{u0} = 0$. Z was 7.496 which was statistically significant at p<0.001. There was strong evidence that between trust variance was not zero.

Table 5.1-7: Null Model, Emergency CS

Fixed Coefficients (261,468 cases in 143 trusts)

						95% Confidence Interval for Odds Ratio	
Model Term	Coefficient	Std. Error	T	Sig.	Odds Ratio	Lower	Upper
Intercept	-1.441	.018	-82.306	<.001	.237	.229	.245

Random Effect

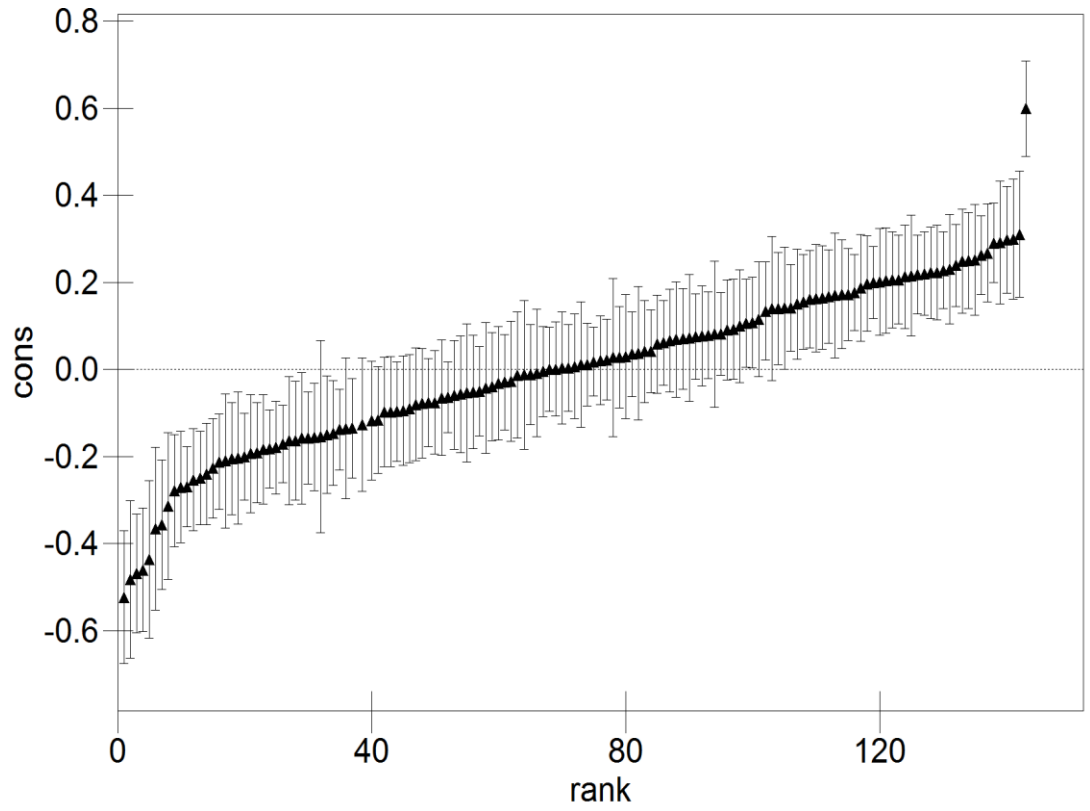
					95% Confidence Interval	
Random Effect Covariance	Estimate	Std. Error	Z	Sig.	Lower	Upper
Var (Intercept)	.040	.005	7.496	<.001	.031	.052

Covariance Structure: Variance components
Subject Specification: Trust ID

The variance partition coefficient (VPC) was calculated as 0.040/(0.040+3.29)=0.012. Thus 1.2% of the residual variation in the emergency CS was attributable to unobserved trusts characteristics.

Using the null model results one can examine the estimates of the trusts effects (residuals $\widehat{\boldsymbol{u}}_{\boldsymbol{0j}}$) by plotting them ranked with their 95% confidence intervals, i.e. creating a 'caterpillar plot' (Figure 7). The caterpillar plot was done in MLwiN 2.26, after applying PQL2 procedure. The residual variation in MLwiN 2.26 using PQL2 was 0.040 (SE 0.005) and the intercept was -1.443 (SE 0.018), which were the similar using GLMM in IBM SPSS 22 PQL procedure (Table 5.1-7).

Figure 7: Caterpillar plot of estimated residuals in 143 trusts for Emergency CS (Null Model)



The caterpillar plot (Figure 7) shows the estimated residuals for the 143 trusts (the outlier trust, top of the graph, was RNH, which had 31.1% emergency CS). It seemed that for more than half of the trusts, the 95% confidence intervals did not overlap the horizontal line at zero, meaning that the incidence of emergency CS in these trusts was significantly above average (for the 42 trusts above the zero line) or below average (for the 35 trusts below the zero line). Where the confidence intervals were wide, this indicated a smaller sample size (fewer women aged 15-44, who were nulliparous, at term, with a singleton, live birth who have experienced emergency CS) within a particular trust, leading to larger standard errors for the estimated trusts residuals $\widehat{\boldsymbol{u}}_{\boldsymbol{0j}}$.

Q2: Do women's socio-demographic characteristics and clinical risk affect their probability of experiencing emergency CS?

To answer this question and in order to examine the strength of the relationships between Level-1 covariates and emergency CS, the two-level random intercept model (the Null) was extended by adding individual-level covariates: maternal age (Ref: 15-19); ethnicity (Ref: white); IMD in quintiles (Ref: least deprived); NICE any risk (Ref: low risk), gestational age (Ref: 37-41); and infant birth weight (Ref: 2500-4500g). The variables were added in three blocks:

- First - maternal age, ethnicity and IMD;
- Second - NICE risk;
- Third - gestational age and infant birth weight.

Table 5.1-8 shows that the fixed effects of all Level-1 predictors were statistically significant at p<.001. This suggests that all of them were potentially important predictors of emergency CS. The effect size of the NICE risk variable was the highest ($F_{1,194021}$ = 2267.4, p < .001), followed by birth weight ($F_{2,194021}$ = 321.8, p < .001), maternal age ($F_{5,194021}$ = 291.2, p < .001), gestational age ($F_{1,194021}$ = 273.1, p < .001) and ethnicity ($F_{3,194021}$ = 106.6, p < .001). The overall effect size of IMD, though significant, was somewhat weaker in explaining emergency CS.

Table 5.1-8: Fixed Effects, Level-1 predictors, Emergency CS

N of cases 194,038; N of trusts 132

Source	F	Nom df1	Denom df2	Sig.
Corrected Model	371.994	16	194021	<.001
Maternal Age	291.226	5	194021	<.001
Ethnicity	106.573	3	194021	<.001
IMD	5.810	4	194021	<.001
NICE Risk	2267.396	1	194021	<.001
Gestational Age	273.074	1	194021	<.001
Birth Weight	321.840	2	194021	<.001

Fixed effects were obtained through Satterthwaite approximation to adjust for denominator degrees of freedom.[105]

Table 5.1-9 presents the estimates of the individual parameters, their standard errors, the odds ratios and their confidence intervals.

The model results showed that the probability of experiencing emergency CS increased significantly with age, adjusting for the effects of other predictors. Older women were significantly more likely to experience emergency CS than the youngest age group (15-19) (within the same trust, and if the same ethnicity, deprivation quintile, clinical risk group, gestation and baby's birth weight). For example women aged 40-44 had a 192% increase in the odds of emergency CS compared to the youngest women aged 15-19 years. The odds of experiencing emergency CS also increased with the level of deprivation, although this was of a smaller magnitude, with most deprived women having an 11% increase in the odds compared to the least deprived. There was a 63% increase in the odds of emergency CS for Black Caribbean women compared to a White women and 132% increase in the odds of emergency CS for high risk women compare to low risk women. The estimated odds of emergency CS for a woman with gestational age greater than 41 weeks were 64% ($e^{.493}$= 1.638) higher than the odds of emergency CS for a woman with gestational age between 37 and 41 weeks. Women with babies weighing more than 4500g had a 216% increase in the odds of emergency CS compared to women with babies weighing between 2500g and 4500g.

[105] Satterthwaite approximation was recommended in IBM SPSS 22, GLMM (model option DF_METHOD=SATTERTHWAITE) for smaller sample size, unbalanced data or complicated covariance structure such as unstructured.

Table 5.1-9: Fixed Coefficients, Level-1 predictors, Emergency CS

N of cases 194,038; N of trusts 132

Model Term	Coeff.	Std. Error	t	Sig.	Odds Ratio	95% Confidence Interval for Odds Ratio Lower	Upper
Intercept	-2.606	.039	-67.137	<.001	.074	.068	.080
Maternal Age							
40-44	1.071	.050	21.341	<.001	2.918	2.644	3.219
35-39	1.006	.037	27.314	<.001	2.733	2.543	2.938
30-34	.775	.031	25.033	<.001	2.171	2.043	2.307
25-29	.571	.032	17.636	<.001	1.771	1.662	1.887
20-24	.313	.032	9.864	<.001	1.368	1.286	1.456
15-19	0[b]	.	.	.	1.000	.	.
Ethnicity							
Mixed/Chinese/Other	.142	.026	5.459	<.001	1.153	1.095	1.213
Black-Caribbean	.490	.029	16.899	<.001	1.633	1.542	1.728
Asian	.236	.026	8.954	<.001	1.266	1.202	1.333
White	0[b]	.	.	.	1.000	.	.
IMD in quintiles							
5 Most Deprived	.106	.023	4.532	<.001	1.112	1.062	1.164
4	.079	.021	3.784	<.001	1.082	1.039	1.127
3	.063	.018	3.459	.001	1.065	1.028	1.104
2	.058	.020	2.888	.004	1.060	1.019	1.103
1 Least Deprived	0[b]	.	.	.	1.000	.	.
NICE Risk							
NICE High Risk	.843	.018	47.617	<.001	2.324	2.245	2.406
NICE Low Risk	0[b]	.	.	.	1.000	.	.
Gestational Age							
>= 42 weeks	.493	.030	16.525	<.001	1.638	1.545	1.737
37-41 weeks	0[b]	.	.	.	1.000	.	.
Birth Weight							
> 4500g	1.151	.047	24.280	<.001	3.161	2.881	3.469
< 2500g	.202	.035	5.810	<.001	1.223	1.143	1.309
2500g-4500g	0[b]	.	.	.	1.000	.	.

b. Reference Categories

The between trust variance was slightly reduced with the additions of age, ethnicity and IMD variables (from 0.040 to 0.033,Table 5.1-12) to the Null model, suggesting

that the distribution of age, ethnicity or IMD varied across trusts; and increased with the addition of NICE risk variable (from 0.033 to 0.042), which stayed the same after adding gestational age and birth weight.

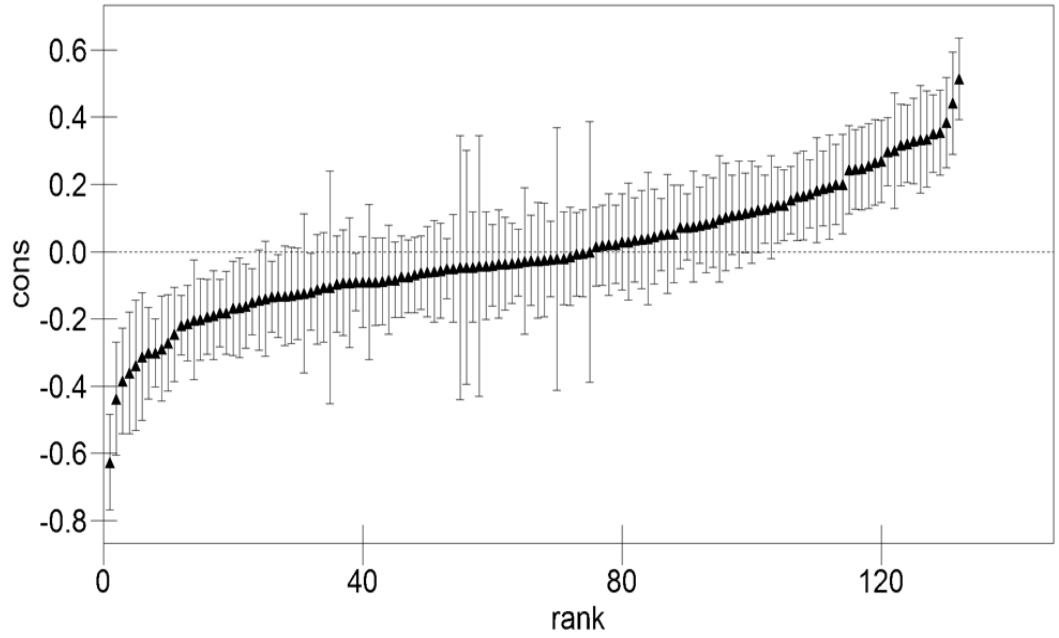
Overall the variance after adding all Level-1 predictors increased to 0.042 and was still significantly different from zero (the Wald Z = 6.990, p<.001). The new variance partition coefficient (VPC) after adding all Level-1 predictors increased from 1.2% to 1.3% ((0.042/(0.042+3.29)=0.013)). Thus 1.3% of the residual variation in the emergency CS was attributable to unobserved trusts characteristics.

The model so far dropped 67,430 cases from 11 trusts (from 261,468 cases in 143 trusts to 194,038 cases in 132 trusts; or 25.8% of the sample), as a result of missing data in all the predictors, but mainly due to missing data in pregnancy gestation at time of birth and birth weight.

Using the risk adjusted model results, the estimates of the trusts effects (residuals $\widehat{\boldsymbol{u}}_{\boldsymbol{0j}}$) were plotted and ranked again with their 95% confidence intervals, i.e. creating a 'caterpillar plot' (Figure 8). The caterpillar plot was produced in MLwiN 2.26.

The caterpillar plot (Figure 8) shows the estimated residuals for the remaining 132 trusts, after risk adjustment. Compared to the Null model (Figure 7), the 95% confidence intervals of the risk adjusted trusts effect estimates, did not overlap the horizontal line at zero for a smaller proportion of trusts (43% of the trusts). This proportion though was still relatively high - the incidence of emergency CS in 23% (30 trusts) of the trusts was significantly above average, and for 20% (27 trusts) it was below the average. The confidence intervals for 9 trusts were quite wide. This indicated a smaller sample size (fewer women aged 15-44, who were nulliparous, at term, with a singleton, live birth) who have had an emergency CS within these 9 trusts, leading to larger standard errors for the estimated trusts residuals $\widehat{\boldsymbol{u}}_{\boldsymbol{0j}}$.

Figure 8: Caterpillar plot of estimated residuals in 132 trusts for Emergency CS, after risk adjustment (Level-1 predictors)



Q3: Do trust configuration and staffing levels influence the probability of emergency CS and do they explain any of the variations in emergency CS between trusts?

To answer this question and to explore the contextual effects on emergency CS, Level-2 (trust level) predictors were added to the random intercept and individual-level covariates model in two blocks.

- The first block included: trust type (London trust; foundation trust; university hospital) and trust configuration - OU/AMU/FMU;
- The second block: the staffing ratios of Consultants O&G; Doctors, Midwives and HCAs all FTE per delivery.

In addition the aim was also to find out if Level-2 predictors explain any of the Level-2 variations in emergency CS between trusts.

The final model takes into account the differences between trusts and tries to explain these differences in terms of staff and trust characteristics. The intercepts from Level-1 ($\boldsymbol{\beta_{0j}}$) were modelled as function of trust type, trust configuration, trusts' staff/birth ratios and a random effect u_{0j}. Modelled in such a way the within trusts intercepts of each trust vary systematically with trust staffing, trust type and trust configuration. Staffing is a trust level variable as the quantity of staff FTE provided to each woman in labour or indeed which staff group assisted at birth were unknown. The coefficients of the other Level-1 predictors were modelled as a fixed slope without random effect. These slopes were not related to the staffing and trust characteristics. If a hypothesis existed that the relationships between women characteristics (clinical risk) and emergency CS differ by trust characteristics (staffing levels), cross level interaction effects should be included in the model (i.e. Staff*NICE Risk for example), which was indeed tested with Q4.

It is worth considering that trusts with fewer women will have a smaller influence on the results but variations in trust size do not pose a problem for multilevel modelling.

Table 5.1-10 presents the fixed effects results of the final model – i.e. a random intercept model, with Level-1 and Level-2 predictors representing the fixed part of the model (the rest of the model (U_{0j}) is the random part that only consists of a random effect). Table 5.1-10 shows the fixed effects for the overall model and individual effects. The model and all fixed effects of Level-1 predictors were highly significant at p<.001 (i.e., the Null hypothesis that the coefficients are equal to 0 can be rejected), of which NICE clinical risk retained primary position in explaining emergency CS ($F_{1,194011}$ = 2272.0, p < .001). In contrast none of the fixed effects of Level-2 predictors (trust type, trust configuration or staffing FTE/birth ratios) were statistically significant at 5%.

Table 5.1-10: Fixed Effects, Level-1 & Level-2 predictors, Emergency CS

N of cases 194,038; N of trusts 132

Source	F	Nom df1	Denom df2	Sig.
Corrected Model	256.330	26	222	<.001
Maternal Age	294.543	5	194011	<.001
Ethnicity	108.551	3	194011	<.001
IMD	5.923	4	194011	<.001
NICE Risk	2271.982	1	194011	<.001
Gestational Age	273.588	1	194011	<.001
Birth Weight	321.981	2	194011	<.001
University Hospital	.099	1	119	.754
London Trust	1.199	1	84	.277
Foundation Trust , Sept 2010	.155	1	86	.695
Trust OU/AMU/FMU	.238	3	74	.869
ZConsultants FTE/birth	.841	1	103	.361
ZDoctors FTE/birth	3.341	1	91	.071
ZMidwives FTE/birth	1.667	1	58	.202
ZHCA FTE/birth	.194	1	71	.661

*OU (Obstetric Unit); AMU (Alongside Midwifery Unit); FMU (Freestanding Midwifery Unit)

Table 5.1-11 presents the estimates of the fixed effects coefficients. The coefficients show the relationship of each model parameter to emergency CS. The expected log-odds of emergency CS for an average woman (aged 15-19, low risk, White, least deprived, with gestational age 37-41 weeks and a baby weighing between 2500g and 4500 g) were -2.603 (SE 0.050). These log-odds correspond to a probability of 0.069=$[e^{-2.603}/(1 + e^{-2.603})]$.

There was a positive association but not statistically significant, between consultant obstetricians FTE/birth and emergency CS and a negative association of doctors, midwives and HCA all FTE/birth with emergency CS. Only the coefficient for doctors FTE/birth approached statistical significance at 10%.

None of the other trust characteristics were statistically significant at 5%. Giving birth in a London trust was positively related to emergency CS; while giving birth in

a University Hospital or Foundation Trust were negatively related to emergency CS (none of them stat significant at 5%).

Table 5.1-11: Fixed Coefficients, Level-1 & Level-2 predictors, Emergency CS

N of cases 194,038; N of trusts 132

Model Term	Coefficient	Std. Error	t	Sig.	Odds Ratio	95% Confidence Interval for Odds Ratio Lower	Upper
Intercept	-2.603	.050	-52.476	<.001	.074	.067	.082
Level-1 predictors							
Maternal Age							
40-44	1.070	.050	21.353	<.001	2.915	2.642	3.216
35-39	1.004	.037	27.371	<.001	2.730	2.541	2.934
30-34	.774	.031	25.068	<.001	2.169	2.042	2.305
25-29	.571	.032	17.647	<.001	1.770	1.661	1.886
20-24	.313	.032	9.856	<.001	1.368	1.285	1.456
15-19	0[b]	.	.	.	1.000	.	.
Ethnicity							
Mixed/Chinese/Other	.140	.026	5.466	<.001	1.150	1.094	1.209
Black-Caribbean	.487	.029	17.099	<.001	1.628	1.540	1.722
Asian	.233	.026	8.869	<.001	1.263	1.199	1.329
White	0[b]	.	.	.	1.000	.	.
IMD in quintiles							
5 Most Deprived	.107	.023	4.598	<.001	1.113	1.063	1.165
4	.079	.021	3.775	<.001	1.082	1.039	1.128
3	.064	.018	3.486	<.001	1.066	1.028	1.105
2	.059	.020	2.908	.004	1.061	1.019	1.103
1 Least Deprived	0[b]	.	.	.	1.000	.	.
NICE Risk							
NICE High Risk	.844	.018	47.665	<.001	2.325	2.246	2.407
NICE Low Risk	0[b]	.	.	.	1.000	.	.
Gestational Age							
>=42 weeks	.493	.030	16.540	<.001	1.638	1.545	1.737
37-41 weeks	0[b]	.	.	.	1.000	.	.
Birth Weight							
>4500g	1.151	.047	24.284	<.001	3.161	2.881	3.469
<2500g	.202	.035	5.819	<.001	1.223	1.143	1.309
2500g-4500g	0[b]	.	.	.	1.000	.	.

Level-2 predictors							
University Hospital	-.015	.047	-.315	.754	.985	.897	1.082
Not Uni-Hospital	0[b]	.	.	.	1.000	.	.
London Trust	.061	.055	1.095	.277	1.063	.952	1.186
Not London Trust	0[b]	.	.	.	1.000	.	.
Foundation Trust	-.015	.039	-.393	.695	.985	.911	1.064
Not Foundation Trust	0[b]	.	.	.	1.000	.	.
OU/AMU/FMU*	-.027	.058	-.459	.648	.974	.867	1.093
OU/FMU	.030	.054	.553	.583	1.030	.925	1.147
OU/AMU	-.007	.047	-.138	.891	.993	.904	1.092
OU only	0[b]	.	.	.	1.000	.	.
ZConsultantsFTE/birth	.022	.024	.917	.361	1.022	.975	1.072
ZDoctors FTE/birth	-.040	.022	-1.828	.071	.961	.921	1.003
ZMidwives FTE/birth	-.028	.021	-1.291	.202	.973	.932	1.015
ZHCA FTE/birth	-.008	.018	-.441	.661	.992	.956	1.029

*OU (Obstetric Unit); AMU (Alongside Midwifery Unit); FMU (Freestanding Midwifery Unit)
b. Reference Categories;

Overall the final model indicates that women characteristics had a strong and significant relationship with emergency CS, particularly NICE clinical risk. The inclusion of Level-2 predictors did not change much the coefficients of Level-1 predictors.

The variance after adding all Level-2 predictors, increased slightly to 0.044 and was significantly different from zero (the Wald Z = 6.767, p<.001, i.e. strong evidence that the between trusts variance was not zero at 1% level). The new variance partition coefficient (VPC) after adding all Level-2 predictors remained at 1.3% ((0.044/(0.044+3.29)=0.013)). Thus a slightly higher but negligible proportion (1.3% compared to 1.2% for the Null model) of the total variation in emergency CS was attributable to between trusts variation after adding all Level-1 and Level-2 predictors to the Null model. Table 5.1-12 presents the changes in the variance at the different modelling stages for emergency CS. It seemed that inclusion of NICE Risk and trust variables increased slightly the between trusts variations in emergency CS.

Table 5.1-12: Intercept Variances from different stages of modelling Emergency CS

					95% Confidence Interval	
Random Effect Covariance	Estimate	Std. Error	Z	Sig.	Lower	Upper
Null	.040	.005	7.496	<.001	.031	.052
Age, Ethnicity, IMD	.033	.005	7.273	<.001	.025	.043
NICE Risk	.042	.006	7.455	<.001	.032	.055
Gestation & Birth Weight	.042	.006	6.990	<.001	.032	.056
Trust Characteristics	.044	.006	6.860	<.001	.033	.058
ZStaff FTE/birth ratios)	.044	.006	6.767	<.001	.033	.058

Covariance Structure: Variance components
Subject Specification: Trust ID

5.1.4 INSTRUMENTAL DELIVERY - MULTILEVEL MODEL

The modelling of instrumental delivery followed the same stages as for emergency CS, and had the same predictors, to allow comparison between models. The modelling stages were determined by the same research questions.

Q1: What is the extent (if any) of between-trusts variation in instrumental delivery?

A Null model was fitted first, e.g. a two-level random intercept model with no predictors to explore the extent (if any) of between-trusts variations in instrumental delivery.

The PQL estimates of the Null model (Table 5.1-13) for instrumental delivery showed that the log-odds of having instrumental delivery in an average trust (one with u_{0j}=0) was estimated at: -1.303 (SE 0.021). The intercept for trust j was: -1.303+ u_{0j} and the variance σ^2_{u0} of u_{0j} was estimated at: 0.059 (SE 0.007). Wald Z was 7.872 which was statistically significant at p<0.001, and provided a strong evidence that between trust variance was not zero.

Table 5.1-13: Null Model, Instrumental Delivery

Fixed Coefficients (261,468 cases in 143 trusts)

						95% Confidence Interval for Odds Ratio	
Model Term	Coefficient	Std. Error	t	Sig.	Odds Ratio	Lower	Upper
Intercept	-1.303	.021	-62.381	<.001	.272	.261	.283

Random Effect

					95% Confidence Interval	
Random Effect Covariance	Estimate	Std. Error	Z	Sig.	Lower	Upper
Var (Intercept)	.059	.007	7.872	<.001	.046	.076

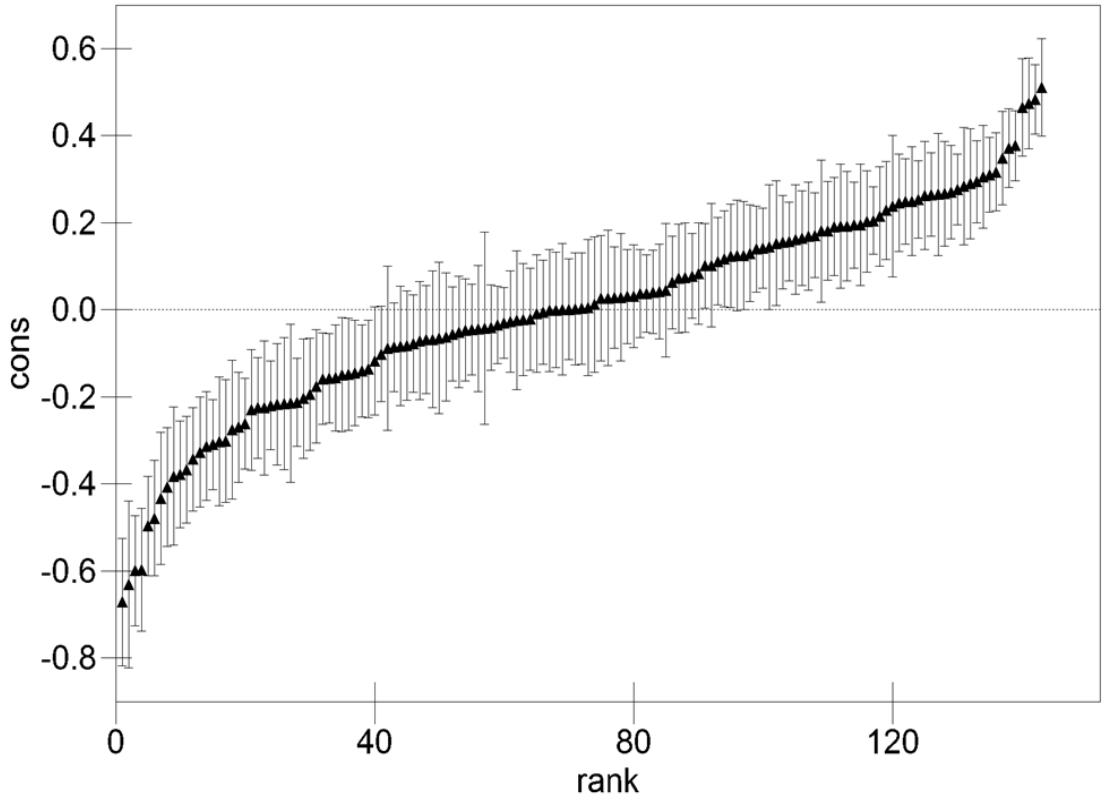
Covariance Structure: Variance components
Subject Specification: Trust ID

Overall there was a slightly higher residual variation between trusts for instrumental delivery (0.059, SE 0.007) compared to emergency CS (0.040, SE 0.005). The variance partition coefficient (VPC) for instrumental delivery was calculated as 0.059/(0.059+3.29)=0.018. Thus 1.8% of the residual variation in instrumental delivery was attributable to unobserved trusts characteristics.

The caterpillar plot (Figure 9) shows the estimated residuals for the143 trusts. It seemed that for 61% of the trusts, the 95% confidence intervals did not overlap the horizontal line at zero, meaning that the incidence of instrumental delivery in these trusts was significantly above average (34% or 48 trusts were above the zero line) or below the average (27% or 39 trusts were below the zero line).

The caterpillar plot was produced in MLwiN 2.26. The residual variation in MLwiN 2.26 using PQL2 was 0.059 (SE 0.007) and the intercept was -1.304 (SE 0.021), which were the same using GLMM in IBM SPSS 22 (Table 5.1-13).

Figure 9: Caterpillar plot of estimated residuals in 143 trusts for Instrumental Delivery (Null Model)



Q2: Do women socio-demographic characteristics and clinical risk affect their probability of experiencing instrumental delivery?

To examine the strength of the relationships between Level-1 covariates and instrumental delivery, the two-level random intercept model (the Null) was extended by adding individual-level covariates in three blocks as before in the model for emergency CS:

- First - maternal age, ethnicity and IMD;
- Second - NICE risk;
- Third - gestational age and infant birth weight.

Table 5.1-14 shows that the fixed effects of all Level-1 predictors were statistically significant at p<.001, which suggested that all of them were potentially important

predictors for having an instrumental delivery. NICE risk did not have the highest effect size among the predictors as with emergency CS, indeed it had one of the lowest. Gestational age had the strongest effect ($F_{1,194021}$ = 172.5, p < .001), followed by maternal age ($F_{5,194021}$ = 167.3, p < .001) and ethnicity ($F_{3,194021}$ = 141.7, p < .001). Overall effect sizes of IMD, NICE risk and birth weight were weaker in predicting instrumental delivery.

Table 5.1-14: Fixed Effects, Level-1 predictors, Instrumental Delivery

N of cases 194,038; N of trusts 132

Source	F	Nom df1	Denom df2	Sig.
Corrected Model	131.623	16	194021	<.001
Maternal Age	167.308	5	194021	<.001
Ethnicity	141.718	3	194021	<.001
IMD	9.968	4	194021	<.001
NICE Risk	19.203	1	194021	<.001
Gestational Age	172.448	1	194021	<.001
Birth Weight	47.824	2	194021	<.001

Table 5.1-15 presents the individual fixed parameter estimates, their standard errors, the odds ratios and confidence intervals. There was a positive curvilinear (inverted-U) relationship between age and instrumental delivery. Women aged 30-34 were more likely to experience instrumental delivery compared to the youngest age group (15-19) (within the same trust, and within the same ethnicity, deprivation quintile, clinical risk group, gestational age and baby's birth weight). They had an 84% increase in odds of instrumental delivery compared to the youngest age group (15-19 years). The odds of experiencing instrumental delivery decreased with the level of deprivation, with most deprived women having a 12% decrease in the odds compared to the least deprived and adjusting for the effects of other predictors (only 4th and 5th quintiles were statistically significant at p<.001). Asian women had slightly higher odds (5%) than white women of instrumental delivery; while Black-Caribbean women had lower odds than white women with a 74% decrease in the odds of

instrumental delivery. The odds of instrumental delivery for high risk women were 8% lower compared to low risk women. There was a positive relationship between gestational age and instrumental delivery (an increase of 32% in the odds of instrumental delivery for women with gestational age greater than 42 weeks compared to 37-41 weeks) and a negative one between infant birth weight and instrumental delivery (the estimated odds of instrumental delivery for women with babies weighing more than 4500g and with babies weighing less than 2500 g were lower by 11% and 32% respectively, both compared to women with babies weighing between 2500g and 4500 g).

Table 5.1-15: Fixed Coefficients, Level-1 predictors, Instrumental Delivery

N of cases 194,038; N of trusts 132

						95% Confidence Interval for Odds Ratio	
Model Term	Coefficient	Std. Error	t	Sig.	Odds Ratio	Lower	Upper
Intercept	-1.619	.034	-48.266	<.001	.198	.185	.212
Maternal Age							
40-44	.363	.042	8.687	<.001	1.438	1.325	1.560
35-39	.561	.029	19.046	<.001	1.752	1.654	1.856
30-34	.609	.024	25.678	<.001	1.839	1.755	1.926
25-29	.511	.021	24.536	<.001	1.667	1.600	1.736
20-24	.280	.022	12.832	<.001	1.323	1.268	1.381
15-19	0[b]	.	.	.	1.000	.	.
Ethnicity							
Mixed/Chinese/Other	-.143	.027	-5.375	<.001	.867	.823	.913
Black-Caribbean	-.768	.041	-18.658	<.001	.464	.428	.503
Asian	.052	.022	2.417	.016	1.054	1.010	1.099
White	0[b]	.	.	.	1.000	.	.
IMD in quintiles							
5 Most Deprived	-.131	.022	-5.996	<.001	.877	.841	.916
4	-.082	.021	-3.914	<.001	.921	.884	.960
3	-.033	.018	-1.826	.068	.967	.933	1.002
2	-.031	.018	-1.783	.075	.969	.936	1.003
1 Least Deprived	0[b]	.	.	.	1.000	.	.
NICE Risk							
NICE High Risk	-.080	.018	-4.382	<.001	.923	.890	.957

NICE Low Risk	0[b]	.	.	.	1.000	.	.
Gestational Age							
>= 42 weeks	.278	.021	13.132	<.001	1.320	1.267	1.376
37-41 weeks	0[b]	.	.	.	1.000	.	.
Birth Weight							
> 4500g	-.114	.047	-2.406	.016	.892	.813	.979
< 2500g	-.382	.041	-9.378	<.001	.682	.630	.739
2500g-4500g	0[b]	.	.	.	1.000	.	.

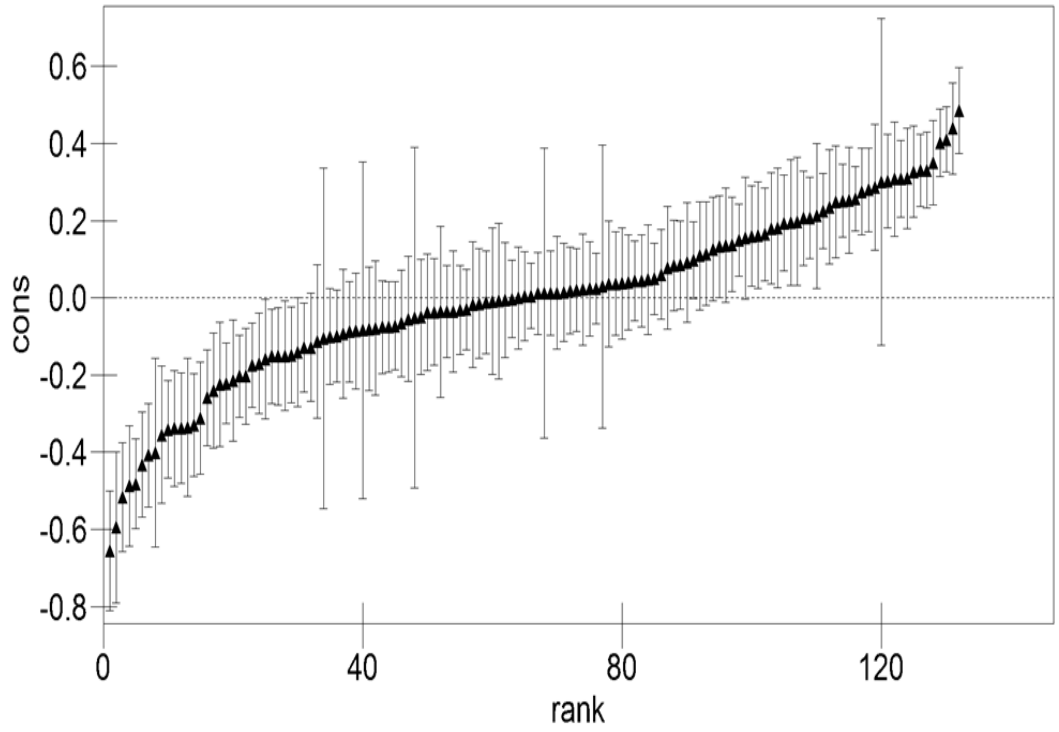
b. Reference categories

This model also dropped 11 trusts, as a result of missing data in all the predictors, but mainly due to missing data in pregnancy gestation at time of birth and birth weight.

Using the risk adjusted model results, the estimates of the trusts effects (residuals $\widehat{\boldsymbol{u}}_{\boldsymbol{0j}}$) were plotted and ranked again with their 95% confidence intervals, i.e. creating a 'caterpillar plot' (Figure 10). The caterpillar plot was produced in MLwiN 2.26.

The caterpillar plot (Figure 10) shows the estimated residuals for the remaining 132 trusts, after risk adjustment. Compared to the Null model (Figure 9), the 95% confidence intervals of the risk adjusted trusts effect estimates, did not overlap the horizontal line at zero for a smaller proportion of trusts (49% vs 61%). This proportion though was still relatively high - the incidence of instrumental delivery in 26% (34 trusts) of the trusts was significantly above average, and for 23% (31 trusts) it was below the average. The confidence intervals for 6 trusts were very wide. This indicated a smaller sample size (fewer women aged 15-44, who were nulliparous, at term, with a singleton, live birth) who have had an instrumental delivery within these 6 trusts, leading to larger standard errors for the estimated trusts residuals $\widehat{\boldsymbol{u}}_{\boldsymbol{0j}}$.

Figure 10: Caterpillar plot of estimated residuals in 132 trusts for Instrumental Delivery, after risk adjustment (Level-1 predictors)



Q3: Do trust configuration and staffing levels influence the probability of instrumental delivery and do they explain any of the variations in the outcome between trusts?

The contextual effects on the instrumental delivery were explored by adding Level-2 (Trust level) predictors to the random intercept and Level-1 predictors in two blocks.

- The first block included: trust type (London trust; foundation trust; university hospital) and trust configuration - OU/AMU/FMU;
- The second block: the staffing ratios of consultant obstetricians; doctors, midwives and HCAs, all standardized ratios of FTE per delivery.

In addition the aim was also to find out if Level-2 predictors explain any of the Level-2 variations in instrumental delivery between trusts.

Table 5.1-16 presents the fixed effects results of the final model – i.e. a random intercept model, with Level-1 and Level-2 predictors representing the fixed part of the model (the rest of the model (U_{0j}) was the random part that only consists of a random effect). Table 5.1-16 shows the fixed effects for the overall model and individual effects. The fixed effects for the overall model and all Level-1 predictors were highly significant at p<.001 (i.e., we can reject the Null hypothesis that the coefficients were equal to 0), of which gestational age, maternal age and ethnicity retained primary positions in predicting instrumental delivery (respectively $F_{1,194011}$ = 171.4; $F_{5,194011}$ = 166.4; $F_{3,194011}$ = 141.0, all p < .001). None of the trusts characteristics (trust type, trust configuration) fixed effects were significant, while standardized consultants FTE/birth ratio was statistically significant at 5% and the fixed effect of the standardized midwives FTE/birth ratio approached significance at 10%.

Table 5.1-16: Fixed Effects, Level-1 & Level-2 predictors, Instrumental Delivery

N of cases 194,038; N of trusts 132

Source	F	Nom df1	Denom df2	Sig.
Corrected Model	85.755	26	282	<.001
Maternal Age	166.353	5	194011	<.001
Ethnicity	141.015	3	194011	<.001
IMD	10.031	4	194011	<.001
NICE Risk	19.130	1	194011	<.001
Gestational Age	171.368	1	194011	<.001
Birth Weight	47.742	2	194011	<.001
University Hospital	.386	1	145	.535
London Trust	.119	1	66	.731
Foundation Trust , Sept 2010	.468	1	74	.496
Trust OU/AMU/FMU	.268	3	110	.848
ZConsultants FTE/birth	9.380	1	72	.003
ZDoctors FTE/birth	.097	1	81	.756
ZMidwives FTE/birth	3.189	1	129	.076
ZHCA FTE/birth	.022	1	81	.882

*OU (Obstetric Unit); AMU (Alongside Midwifery Unit); FMU (Freestanding Midwifery Unit)

The final model (Table 5.1-17) indicates that women characteristics had a strong and significant relationship with instrumental delivery (exception 2nd/3rd IMD quintiles). The inclusion of Level-2 predictors did not change much the coefficients of Level-1 predictors. One SD increase in the midwives FTE/birth ratio decreased the log-odds of instrumental delivery by -.051, ($p<0.10$). One SD increase in consultants FTE/birth ratio increased the odds of instrumental delivery by 7.6% (OR 1.076; CI 1.026-1.129). None of the other Level-2 Trust and staff variables were statistically significant at 5%.

Table 5.1-17: Fixed Coefficients, Level-1 & Level-2, Instrumental Delivery

N of cases 194,038; N of trusts 132

						95% Confidence Interval for Odds Ratio	
Model Term	Coefficient	Std. Error	t	Sig.	Odds Ratio	Lower	Upper
Intercept	-1.649	.041	-39.87	<.001	.192	.177	.209
Level-1 predictors							
Maternal Age							
40-44	.362	.042	8.629	<.001	1.436	1.323	1.559
35-39	.560	.030	18.873	<.001	1.750	1.651	1.855
30-34	.608	.024	25.570	<.001	1.837	1.754	1.925
25-29	.511	.021	24.472	<.001	1.666	1.600	1.736
20-24	.280	.022	12.817	<.001	1.323	1.268	1.381
15-19	0[b]	.	.	.	1.000	.	.
Ethnicity							
Mixed/Chinese/Other	-.145	.026	-5.467	<.001	.865	.822	.911
Black-Caribbean	-.770	.041	-18.58	<.001	.463	.427	.502
Asian	.052	.022	2.396	.017	1.053	1.009	1.098
White	0[b]	.	.	.	1.000	.	.
IMD in quintiles							
5 Most Deprived	-.131	.022	-6.008	<.001	.877	.840	.915
4	-.083	.021	-3.935	<.001	.921	.884	.959
3	-.033	.018	-1.820	.069	.967	.933	1.003
2	-.031	.018	-1.795	.073	.969	.936	1.003
1 Least Deprived	0[b]	.	.	.	1.000	.	.
NICE Risk							
NICE High Risk	-.080	.018	-4.374	<.001	.923	.891	.957

NICE Low Risk	0[b]	.	.	.	1.000	.	.
Gestational Age							
>=42 weeks	.277	.021	13.091	<.001	1.320	1.266	1.376
37-41 weeks	0[b]	.	.	.	1.000	.	.
Birth Weight							
>4500g	-.114	.047	-2.408	.016	.892	.813	.979
<2500g	-.382	.041	-9.372	<.001	.683	.630	.739
2500g-4500g	0[b]	.	.	.	1.000	.	.
Level-2 predictors							
University Hospital	.034	.054	.621	.535	1.034	.929	1.152
Not Uni-Hospital	0[b]	.	.	.	1.000	.	.
London Trust	-.020	.057	-.345	.731	.980	.874	1.099
Not London Trust	0[b]	.	.	.	1.000	.	.
Foundation Trust	.028	.041	.684	.496	1.029	.948	1.116
Not Foundation Trust	0[b]	.	.	.	1.000	.	.
OU/AMU/FMU*	-.001	.069	-.013	.989	.999	.872	1.145
OU/FMU	-.022	.068	-.323	.747	.978	.856	1.119
OU/AMU	.041	.054	.765	.446	1.042	.937	1.159
OU only	0[b]	.	.	.	1.000	.	.
ZConsultantsFTE/birth	.074	.024	3.063	.003	1.076	1.026	1.129
ZDoctors FTE/birth	.007	.023	.312	.756	1.007	.962	1.055
ZMidwives FTE/birth	-.051	.028	-1.786	.076	.950	.898	1.006
ZHCA FTE/birth	.003	.021	.149	.882	1.003	.962	1.046

b - Reference categories;
*OU (Obstetric Unit); AMU (Alongside Midwifery Unit); FMU (Freestanding Midwifery Unit)

The variance was reduced after adding age, ethnicity and IMD, increased slightly with the addition of gestational age and trust characteristics, and was reduced again after adding all standardized staffing/birth ratios (Table 5.1-18). There was a strong evidence that the between trusts variance was not zero ($p<.001$) at all the modelling stages (see Wald Z statistics in Table 5.1-18). The final variance partition coefficient (VPC) after adding all Level-1 & Level-2 predictors was slightly reduced to 1.6% (0.053/(0.053+3.29)=0.016), compared to the Null's model VPC of 1.8%. Thus 1.6% of the total variation in instrumental delivery remained attributable to between trusts variation. Table 5.1-18 presents the changes in the variance at the different modelling stages for instrumental delivery.

Table 5.1-18: Intercept Variances from different stages of modelling Instrumental Delivery

					95% Confidence Interval	
Random Effect Covariance	Estimate	Std. Error	Z	Sig.	Lower	Upper
Null	.059	.007	7.872	<.001	.046	.076
Age, Ethnicity, IMD	.054	.007	7.775	<.001	.042	.070
NICE Risk	.054	.007	7.779	<.001	.042	.070
Gestation & Birth Weight	.055	.008	7.158	<.001	.042	.072
Trust Characteristics	.056	.008	6.967	<.001	.042	.074
ZStaff FTE/birth ratios	.053	.008	6.801	<.001	.040	.070

Covariance Structure: Variance components
Subject Specification: Trust ID

5.1.5 NORMAL BIRTH – MULTILEVEL MODEL

The total number for the normal birth sub-sample was 214,920 deliveries in 127 trusts. The NICE risk variable was reversed, so that high risk became the reference category, and the results were interpreted for the low risk group of women who by definition were more likely to experience normal birth.

The multilevel modelling of normal birth followed the same stages as for emergency CS and instrumental delivery, with the same research questions.

Q1: What is the extent (if any) of between-trusts variation in normal birth?

The Null model was fitted first, e.g. two-level random intercept model with no predictors to explore the extent (if any) of between-trusts variations in normal birth.

The PQL estimates of the Null model (Table 5.1-19) for normal birth showed that the log-odds of having normal birth in an average trust (one with u_{0j}=0) was estimated at: -0.826 (SE 0.021). The intercept for trust j was: -0.826+ u_{0j} and the variance σ^2_{u0} of u_{0j} was estimated at: 0.053 (SE 0.007). Wald Z was 7.445 which was statistically significant at p<0.001, and provided a strong evidence that between trust variance was not zero.

Table 5.1-19: Null Model, Normal Birth

Fixed Coefficients (214,920 cases in 127 trusts)

						95% Confidence Interval for Odds Ratio	
Model Term	Coefficient	Std. Error	t	Sig.	Odds Ratio	Lower	Upper
Intercept	-.826	.021	-39.532	<.001	. 438	.420	.456

Random Effect

					95% Confidence Interval	
Random Effect Covariance	Estimate	Std. Error	Z	Sig.	Lower	Upper
Var (Intercept)	.053	.007	7.445	<.001	.040	.068

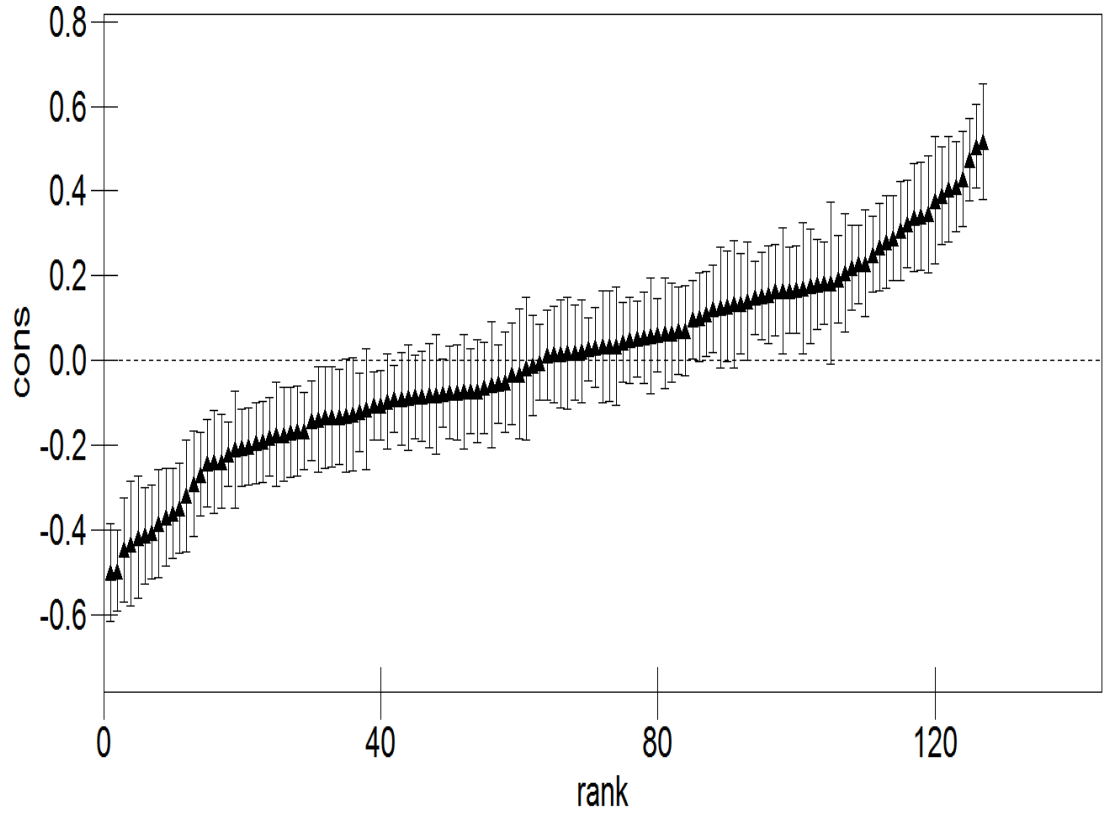
Covariance Structure: Variance components
Subject Specification: Trust ID

Overall the residual variation between trusts in the Null model for normal birth (0.053, SE 0.007) was slightly higher compared to emergency CS (0.040, SE 0.005) but lower compared to instrumental delivery (0.059, SE 0.007). The variance partition coefficient (VPC) for normal birth was calculated as 0.053/(0.053+3.29)=0.016. Thus 1.6% of the residual variation in normal birth was attributable to unobserved trusts characteristics.

The caterpillar plot (Figure 11) shows the estimated residuals for 127 trusts. The Null model of trust effect estimates for normal birth were used in the caterpillar plot, by ranking the residuals ($\widehat{\boldsymbol{u}}_{\boldsymbol{0j}}$) with their 95% confidence intervals (Figure 11). For 58% of the trusts, the 95% confidence intervals did not overlap the horizontal line at zero, meaning that the incidence of normal birth in these trusts was significantly above average (28% or 36 trusts were above the zero line) or below the average (30% or 38 trusts were below the zero line).

The caterpillar plot was produced in MLwiN 2.26, after applying PQL2 procedure. The residual variation in MLwiN 2.26 using PQL2 was 0.052 (SE 0.007) and the intercept was -0.827 (SE 0.021), which were similar using PQL procedure in GLMM of IBM SPSS 22 (Table 5.1-19).

Figure 11: Caterpillar plot of estimated residuals in 127 trusts for Normal Birth (Null Model)



Q2: Do women's socio-demographic characteristics and clinical risk affect their probability of having a normal birth?

The two-level random intercept models (the Null model) was extended by adding individual-level predictors in blocks: first, maternal age; ethnicity and IMD in quintiles; second, NICE risk and finally gestational age and infant birth weight to examine the strength of the relationships between Level-1 predictors and normal birth outcome. Table 5.1-20 shows that the overall model and all individual level covariates were statistically significant at p<.001, suggesting they were all potentially important predictors of normal birth. NICE clinical risk had the highest effect size ($F_{1,175485}$ = 2362.4, p < .001), as in emergency CS model, in predicting

normal birth. Gestational age, maternal age and birth weight had similar size effects ($F_{1,175485}$ = 173.7; $F_{5,175485}$ = 154.0 and $F_{2,175485}$ = 137.8, all p < .001). Ethnicity and IMD were weaker in predicting normal birth (Table 5.1-20).

Table 5.1-20: Fixed Effects, Level-1 predictors, Normal Birth

N of cases 175,502; N of trusts 121

Source	F	Nom df1	Denom df2	Sig.
Corrected Model	302.027	16	175485	<.001
Maternal Age	154.014	5	175485	<.001
Ethnicity	39.533	3	175485	<.001
IMD	5.712	4	175485	<.001
NICE Risk	2362.373	1	175485	<.001
Gestational Age	173.684	1	175485	<.001
Birth Weight	137.813	2	175485	<.001

Table 5.1-21 presents the individual fixed parameter estimates, their standard errors, odds ratios and 95% confidence intervals. There was a negative relationship between age and normal birth. Older women were less likely to experience normal birth compared to the youngest age group (15-19) (within the same trust, and within the same ethnicity group, deprivation quintile, clinical risk group, gestational age and baby's birth weight). Nulliparous women aged 30-34 years had a 50% decrease in the odds of normal birth compared to the youngest age group (15-19 years). The chances of normal birth increased with the level of deprivation, with most deprived women having a 12% increase in the odds compared to the least deprived (2^{nd} and 3^{rd} quintiles were not statistically significant at 5%). Asian women had 17% lower odds, while Black-Caribbean women had 26% higher odds of normal birth both compared to white women. The odds of normal birth for low risk women were nearly 5 times higher compared to low risk women adjusting for the effects of other predictors. There was a negative relationship between gestational age and normal birth (a reduction of 58% in the odds of normal birth for women with gestational age greater than 42 weeks compared to 37-41 weeks); a negative one between birth weight above 4500g and normal birth (64% reduction in the odds of normal birth for women with babies weighing more than 4500g compared to women with babies weighing

between 2500g and 4500 g); and a positive relationship between birth weight of less than 2500g and normal birth (17% increase in the odds).

Table 5.1-21: Fixed Coefficients, Level-1 predictors, Normal Birth

N of cases 175,502; N of trusts 121

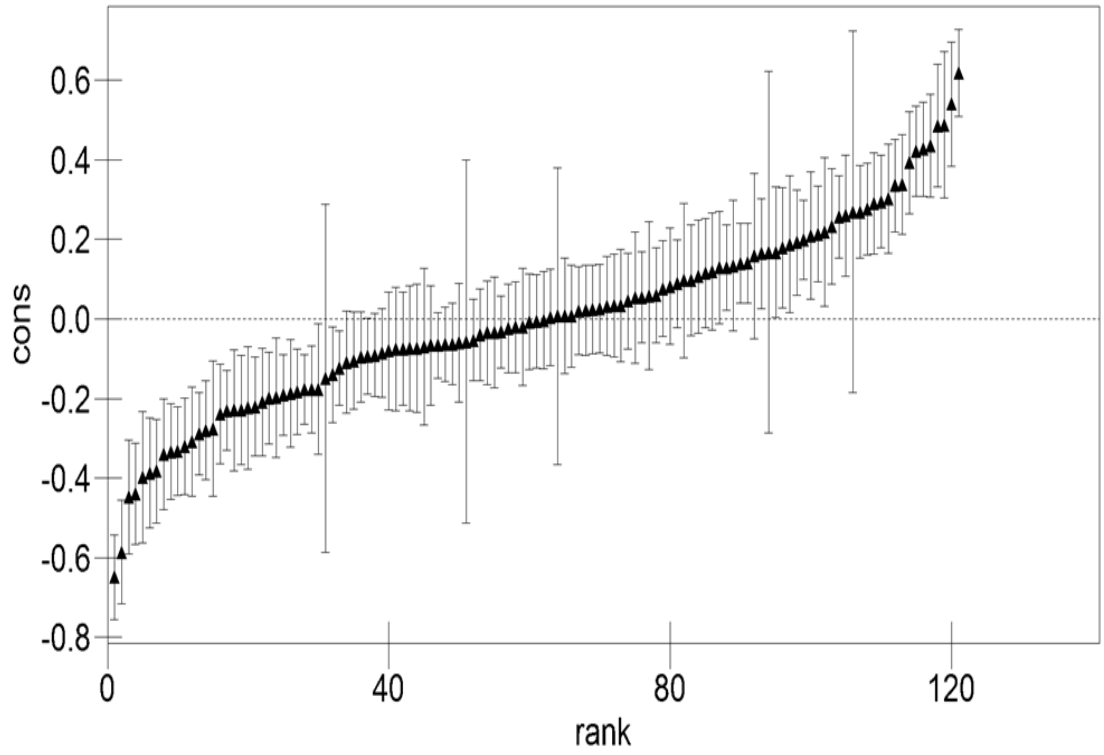
						95% Confidence Interval for Odds Ratio	
Model Term	Coefficient	Std. Error	T	Sig.	Odds Ratio	Lower	Upper
Intercept	-1.423	.047	-30.46	<.001	.241	.220	.264
Maternal Age							
40-44	-.957	.071	-13.43	<.001	.384	.334	.442
35-39	-.840	.040	-21.21	<.001	.432	.399	.467
30-34	-.699	.028	-25.09	<.001	.497	.471	.525
25-29	-.493	.022	-22.35	<.001	.611	.585	.638
20-24	-.232	.022	-10.61	<.001	.793	.760	.828
15-19	0[b]	.	.	.	1.000	.	.
Ethnicity							
Mixed/Chinese/Other	.050	.031	1.629	.103	1.051	.990	1.116
Black-Caribbean	.233	.037	6.249	<.001	1.262	1.173	1.357
Asian	-.184	.024	-7.813	<.001	.832	.795	.871
White	0[b]	.	.	.	1.000	.	.
IMD in quintiles							
5 Most Deprived	.109	.025	4.359	<.001	1.115	1.062	1.171
4	.084	.024	3.458	.001	1.088	1.037	1.141
3	.027	.021	1.330	.184	1.028	.987	1.070
2	.037	.020	1.886	.059	1.038	.999	1.078
1 Least Deprived	0[b]	.	.	.	1.000	.	.
NICE Risk							
NICE Low Risk	1.560	.032	48.604	<.001	4.760	4.470	5.069
NICE High Risk	0[b]	.	.	.	1.000	.	.
Gestational Age							
>= 42 weeks	-.854	.065	-13.18	<.001	.426	.375	.483
37-41 weeks	0[b]	.	.	.	1.000	.	.
Birth Weight							
> 4500g	-1.028	.064	-16.13	<.001	.358	.316	.405
< 2500g	.160	.038	4.184	<.001	1.173	1.089	1.265
2500g-4500g	0[b]	.	.	.	1.000	.	.

b. Reference category

The model for normal birth dropped additional 6 trusts and 18.3% of the observations because of missing data, mainly in gestational age and birth weight variables.

The caterpillar plot (Figure 12) shows the estimated residuals for the remaining 121 trusts, after risk adjustment. The risk adjusted trust effect estimates for normal birth were used in the caterpillar plot, by ranking the residuals ($\widehat{\boldsymbol{u}}_{\boldsymbol{0j}}$) with their 95% confidence intervals (Figure 12). Compared to the Null model (Figure 11), the 95% confidence intervals of the risk adjusted trusts effect estimates, did not overlap the horizontal line at zero for a smaller proportion of trusts (51% vs 58%). This proportion though was still relatively high - the incidence of normal birth in 25% (30 trusts) of the trusts was significantly above average, and for 26% (32 trusts) it was below the average. The confidence intervals for 5 trusts were quite wide. This indicated a smaller sample size (fewer women aged 15-44, who were nulliparous, at term, with a singleton, live birth) who have had a normal birth within these 5 trusts, leading to larger standard errors for the estimated trusts residuals $\widehat{\boldsymbol{u}}_{\boldsymbol{0j}}$.

Figure 12: Caterpillar plot of estimated residuals in 121 trusts for Normal Birth, after risk adjustment (Level-1 predictors)



Q3: Do trusts configuration and staffing levels have an influence on the probability of having a normal birth and do they explain any of the variations in normal birth between trusts?

The contextual effects on normal birth were explored by adding Level-2 (trust level) predictors to the random intercept and Level-1 predictors in two blocks.

- The first block included: trust type (London trust; foundation trust; university hospital) and trust configuration - OU/AMU/FMU;
- The second block: the staffing ratios of consultant obstetricians; doctors, midwives and HCAs, all standardized ratios of FTE per birth.

In addition the aim was also to find out if Level-2 predictors explain any of the Level-2 variations in normal birth between trusts.

Table 5.1-22 presents the fixed effects results of the final model – i.e. a random intercept model, with Level-1 and Level-2 predictors representing the fixed part of the model (the rest of the model (U_{0j}) was the random part that only consists of a random effect). The fixed effects for the overall model and all Level-1 predictors were highly significant at p<.001 (i.e., we can reject the Null hypothesis that the coefficients were equal to 0), of which NICE clinical risk retained primary positions in predicting normal birth (respectively $F_{1,175475}$ = 2347.5, p < .001). None of the trusts characteristics (trust type or configuration) fixed effects were significant at 5% (university hospital approached significance at 10%), while standardized midwives FTE/birth ratio was statistically significant at 5%.

Table 5.1-22: Fixed Effects, Level-1 & Level-2 predictors, Normal Birth

N of cases 175,502; N of trusts 121

Source	F	Nom df1	Denom df2	Sig.
Corrected Model	195.308	26	168	<.001
Maternal Age	154.130	5	175475	<.001
Ethnicity	39.302	3	175475	<.001
IMD	5.887	4	175475	<.001
NICE Risk	2347.477	1	175475	<.001
Gestational Age	173.256	1	175475	<.001
Birth Weight	137.547	2	175475	<.001
University Hospital	2.933	1	134	.089
London Trust	1.488	1	93	.226
Foundation Trust , Sept 2010	.877	1	68	.352
Trust OU/AMU/FMU	1.693	3	74	.176
ZConsultants FTE/birth	.268	1	60	.606
ZDoctors FTE/birth	1.285	1	109	.259
ZMidwives FTE/birth	5.904	1	39	.020
ZHCA FTE/birth	.253	1	37	.618

*OU (Obstetric Unit); AMU (Alongside Midwifery Unit); FMU (Freestanding Midwifery Unit)

The final model (Table 5.1-23) indicated that women's characteristics had a strong and significant relationship (p<.001) with normal birth (exception 2nd/3rd IMD quintiles and ethnicity group for Mixed/Chinese/Other). The inclusion of Level-2

predictors did not alter much the coefficients of Level-1 predictors. One SD increase in midwives FTE/birth ratio increased the odds of normal birth by 5.6% (OR 1.056; CI 1.009-1.105). None of the other Level-2 trust or staff variables were statistically significant at 5%.

Table 5.1-23: Fixed Coefficients, Level-1 & Level-2 predictors, Normal Birth

N of cases 175,502; N of trusts 121

						95% Confidence Interval for Odds Ratio	
Model Term	Coefficient	Std. Error	t	Sig.	Odds Ratio	Lower	Upper
Intercept	-1.403	.057	-24.807	<.001	.246	.220	.275
Level-1 predictors							
Maternal Age							
40-44	-.956	.071	-13.424	<.001	.384	.334	.442
35-39	-.839	.040	-21.181	<.001	.432	.400	.467
30-34	-.698	.028	-25.125	<.001	.498	.471	.525
25-29	-.492	.022	-22.337	<.001	.611	.585	.638
20-24	-.231	.022	-10.599	<.001	.794	.760	.828
15-19	0[b]	.	.	.	1.000	.	.
Ethnicity							
Mixed/Chinese/Other	.053	.030	1.749	.080	1.055	.994	1.119
Black-Caribbean	.237	.037	6.403	<.001	1.268	1.179	1.363
Asian	-.181	.023	-7.711	<.001	.835	.797	.874
White	0[b]	.	.	.	1.000	.	.
IMD in quintiles							
5 Most Deprived	.109	.025	4.423	<.001	1.116	1.063	1.171
4	.085	.025	3.479	.001	1.089	1.038	1.143
3	.027	.021	1.331	.183	1.028	.987	1.070
2	.037	.020	1.899	.058	1.038	.999	1.078
1 Least Deprived	0[b]	.	.	.	1.000	.	.
NICE Risk							
NICE Low Risk	1.560	.032	48.451	<.001	4.761	4.470	5.071
NICE High Risk	0[b]	.	.	.	1.000	.	.
Gestational Age							
>=42 weeks	-.854	.065	-13.163	<.001	.426	.375	.483
37-41 weeks	0[b]	.	.	.	1.000	.	.
Birth Weight							

>4500g	-1.028	.064	-16.112	<.001	.358	.316	.405
<2500g	.160	.038	4.181	<.001	1.173	1.089	1.264
2500g-4500g	0[b]	.	.	.	1.000	.	.
Level-2 predictors							
University Hospital	-.098	.057	-1.713	.089	.907	.809	1.015
Not Uni-Hospital	0[b]	.	.	.	1.000	.	.
London Trust	-.084	.069	-1.220	.226	.919	.801	1.054
Not London Trust	0[b]	.	.	.	1.000	.	.
Foundation Trust	-.040	.043	-.936	.352	.960	.881	1.047
Not Foundation Trust	0[b]	.	.	.	1.000	.	.
OU/AMU/FMU*	.121	.074	1.622	.108	1.128	.973	1.307
OU/FMU	.120	.063	1.887	.064	1.127	.993	1.280
OU/AMU	.045	.055	.825	.412	1.046	.938	1.167
OU only	0[b]	.	.	.	1.000	.	.
ZConsultantsFTE/birth	-.013	.025	-.518	.606	.987	.940	1.037
ZDoctors FTE/birth	-.030	.027	-1.134	.259	.970	.920	1.023
ZMidwives FTE/birth	.055	.022	2.430	.020	1.056	1.009	1.105
ZHCA FTE/birth	-.009	.019	-.503	.618	.991	.954	1.029

b. Reference category.
*OU (Obstetric Unit); AMU (Alongside Midwifery Unit); FMU (Freestanding Midwifery Unit)

The variance after adding all variables, increased slightly from 0.053 to 0.055 and was significantly different from zero (the Wald Z = 6.629, p<.001, i.e. strong evidence that the between trusts variance was not zero at 1% level). The new variance partition coefficient (VPC) after adding all Level-2 predictors remained at 1.6% ((0.055/(0.055+3.29)). Thus 1.6% of the total variation in normal birth was attributable to between trusts variation after adding all Level-1 and Level-2 predictors to the Null model. Table 5.1-24 presents the changes in the variance at the different modelling stages for normal birth. It seemed that inclusion of NICE risk increased between trust variations in normal birth.

Table 5.1-24: Intercept Variances from different stages of modelling Normal Birth

					95% Confidence Interval	
Random Effect Covariance	Estimate	Std. Error	Z	Sig.	Lower	Upper
Null	.053	.007	7.445	<.001	.040	.068
Age, Ethnicity, IMD	.048	.007	7.323	<.001	.036	.062
NICE Risk	.061	.008	7.394	<.001	.047	.079
Gestation & Birth Weight	.060	.009	6.979	<.001	.046	.080
Trust Characteristics	.056	.008	6.772	<.001	.042	.075
ZStaff FTE/birth ratios	.055	.008	6.629	<.001	.041	.074

Covariance Structure: Variance components
Subject Specification: Trust ID

5.2 CONCLUSIONS

In this relatively homogeneous sample of women only around 1%-2% of the residual variations in the outcomes were attributable to unobserved trusts characteristics. In the unadjusted models, the between-trust variances were: 0.040 (SE 0.005) for emergency CS; 0.059 (SE 0.007) for instrumental delivery and 0.053 (SE 0.007) for normal birth. Adjusting for socio-demographics, clinical risk, staffing and trusts characteristics did not reduce substantially the unexplained trust-level variations, in fact inclusion of clinical risk increased the unexplained variations in emergency CS and normal birth; inclusion of trust characteristics increased the variance in emergency CS and instrumental delivery and the inclusion of gestation and birth weight increased the unexplained variations in instrumental delivery. The adjusted models variances were: emergency CS 0.044 (SE 0.007); instrumental delivery 0.053 (SE 0.008); normal birth 0.055 (SE 0.008).

Between trusts and for all women, the standardized consultants FTE/birth ratio was positively related to the probability of instrumental delivery (OR=1.076, CI 1.026-1.129, $p<0.05$), and the standardized midwives FTE/birth ratio was positively related to the probability of normal birth (OR=1.056, CI 1.009-1.105, $p<0.05$). At 10% level of significance, standardized doctors FTE/birth ratio was negatively related to emergency CS (OR=0.961, CI 0.921-1.003, $p<0.10$) and the standardized midwives FTE/birth ratio was negatively related to instrumental delivery (OR=0.950, CI 0.898-1.006, $p<0.10$).

Overall age, clinical risk, birth weight and gestational age had strong relationships with all outcomes. Ethnicity and IMD, though significant were weaker predictors of the three indicators.

Significant associations between certain maternity staff groups and outcomes were identified. Individual level covariates did not substantially reduce the high within-trust variation in outcomes; neither were low between-trust variations affected much by staffing levels and trusts characteristics. The analyses of missing data (see

Chapter 6 for results) did not alter these results. Other unaccounted for factors such as obesity, smoking, organisational culture, differences in practice style and models of care may be able to explain further the variations in outcomes.

6 CHAPTER 6 MODELS EXTENSION

6.1 MISSING DATA ANALYSES

Due to missing data in Level-1 predictors, the final multilevel models for emergency CS and instrumental delivery dropped 25.8% of the observations (11 trusts). 18.3% of the observations (6 trusts) were dropped in the final model for normal birth.

Given these high percentages of data loss, a sensitivity analysis was performed by including the missing data as extra levels in each of the categorical explanatory Level-1 predictors in the final models (with all Level-1 & Level-2 predictors).

6.1.1 EMERGENCY CS – MISSING DATA ANALYSIS

The inclusion of the missing data in the emergency CS model did not substantially change the fixed effect estimates or the direction of the relationships between the categorical predictors and emergency CS outcome. Compared to the emergency CS model which excluded missing data (Table 5.1-8), the only difference in the model with missing data included was the change in the standardized midwives FTE/birth effect, which became statistically significant at 5% ($F_{1,92}$ = 4.327, p < .05, Table 6.1-1).

The results from the multilevel model for emergency CS with missing data included showed, that the fixed coefficients of the missing categories in all the Level-1 categorical variables were not significant at 5%, with the exception for maternal age, which was positive and significant at 1% (Table 6.1-2).

One SD increase in the midwives FTE/birth ratio, reduced the odds of emergency CS by 4% (OR 0.959, CI 0.922-0.998, p<.05) (Table 6.1-2).

The variance in the missing data model for emergency CS remained at 0.044 (SE 0.006) (Table 6.1-3), which was the same in the model of emergency CS with excluded missing data (Table 5.1-12).

6.1.2 INSTRUMENTAL DELIVERY – MISSING DATA ANALYSIS

Results from the missing data model for instrumental delivery were not much different from the model which dropped 25.8% of the observations. None of the missing data coefficients were significant at 5% (ethnicity being an exception, its missing category coefficient was negative and significant at 5%) (Table 6.1-5). The relationship between standardized consultants FTE/birth and instrumental delivery remained positive and statistically significant. The inclusion of missing data slightly reduced the variance to 0.052 (SE 0.007) (Table 6.1-6) from 0.053 (SE 0.008) in the model without missing data (Table 5.1-18).

6.1.3 NORMAL BIRTH – MISSING DATA ANALYSIS

Results from the missing data model for normal birth were similar to the model which dropped 18.3% of the observations. The missing data coefficients for ethnicity and gestational age were positive and significant at 1% and 5% respectively (Table 6.1-8). None of the other coefficients for the missing data were significant at 5% and their inclusion did not affect much the strength or the direction of the relationships between normal birth and the rest of the variables. Standardized midwives FTE/birth remained positively related to normal birth and statistically significant at 5%. The coefficient for the trusts, which had both obstetric and freestanding midwifery units became significant at 5% and was positively related to normal birth. Inclusion of missing data slightly increased the variance to 0.056 (SE 0.008) in (Table 6.1-9), from 0.055 (SE 0.008) in the model without missing data (Table 5.1-24).

Overall the results from the missing data models were very similar to the results from the models which excluded the missing data. Few missing data coefficients were significant at 5% and the inclusion of missing data changed the strength of the relationship between standardized midwives FTE/birth and emergency CS (negative and significant at 5%).

6.1.4 TABLES FROM MISSING DATA SENSITIVITY ANALYSES

6.1.4.1 TABLES FOR EMERGENCY CS MISSING DATA ANALYSIS

Table 6.1-1: Fixed Effects Level-1 & Level-2 predictors, Emergency CS, missing data model

N of cases 261,468; N of trusts 143

Source	F	Nom df1	Denom df2	Sig.
Corrected Model	225.763	31	318	<.001
Maternal Age	275.526	6	261436	<.001
Ethnicity	69.916	4	261436	<.001
IMD	5.592	5	261436	<.001
NICE Risk	2553.638	1	261436	<.001
Gestational Age	155.469	2	47850	<.001
Birth Weight	258.724	3	243784	<.001
University Hospital	.627	1	122	.430
London Trust	.950	1	101	.332
Foundation Trust , Sept 2010	.360	1	100	.550
Trust OU/AMU/FMU	.314	3	81	.815
ZConsultants FTE/birth	1.535	1	116	.218
ZDoctors FTE/birth	2.467	1	124	.119
ZMidwives FTE/birth	4.327	1	92	.040
ZHCA FTE/birth	.054	1	77	.816

Table 6.1-2: Fixed Coefficients, Level-1 & Level-2 predictors, Emergency CS, missing data model

N of cases 261,468; N of trusts 143

Model Term	Coefficient	Std. Error	t	Sig.	Odds Ratio	95% Confidence Interval for Odds Ratio Lower	Upper
Intercept	-2.597	.045	-57.238	<.001	.074	.068	.081
Level-1 predictors							
Maternal Age							
Missing	.456	.105	4.325	<.001	1.578	1.283	1.940
40-44	1.066	.042	25.114	<.001	2.904	2.672	3.156
35-39	.989	.031	31.880	<.001	2.688	2.529	2.856
30-34	.777	.027	28.885	<.001	2.174	2.063	2.292
25-29	.586	.028	21.068	<.001	1.797	1.702	1.898
20-24	.330	.027	12.040	<.001	1.391	1.319	1.468
15-19	0[b]	.	.	.	1.000	.	.
Ethnicity							
Missing	-.008	.025	-.314	.754	.992	.946	1.041
Mixed/Chinese/Other	.122	.023	5.301	<.001	1.130	1.080	1.183
Black-Caribbean	.435	.031	13.969	<.001	1.544	1.453	1.642
Asian	.230	.024	9.535	<.001	1.259	1.200	1.319
White	0[b]	.	.	.	1.000	.	.
IMD in quintiles							
Missing	.103	.072	1.422	.155	1.108	.962	1.276
5 Most Deprived	.092	.020	4.574	<.001	1.097	1.054	1.141
4	.080	.019	4.321	<.001	1.083	1.045	1.123
3	.063	.015	4.124	<.001	1.066	1.034	1.098
2	.044	.017	2.634	.008	1.045	1.011	1.080
1 Least Deprived	0[b]	.	.	.	1.000	.	.
NICE Risk							
High Risk	.856	.017	50.534	<.001	2.355	2.278	2.434
Low Risk	0[b]	.	.	.	1.000	.	.
Gestational Age							
Missing	.008	.041	.205	.838	1.008	.930	1.093
>=42 weeks	.504	.029	17.602	<.001	1.655	1.565	1.750
37-41 weeks	0[b]	.	.	.	1.000	.	.
Birth Weight							
Missing	.063	.052	1.218	.223	1.065	.962	1.178

>4500g	1.137	.043	26.331	<.001	3.117	2.864	3.392
<2500g	.247	.030	8.151	<.001	1.280	1.207	1.359
2500g-4500g	0[b]	.	.	.	1.000	.	.
Level-2 predictors							
University Hospital	-.034	.044	-.792	.430	.966	.886	1.053
Not Uni-Hospital	0[b]	.	.	.	1.000	.	.
London Trust	.053	.054	.975	.332	1.054	.947	1.173
Not London Trust	0[b]	.	.	.	1.000	.	.
Foundation Trust	-.022	.037	-.600	.550	.978	.910	1.052
Not Foundation Trust	0[b]	.	.	.	1.000	.	.
OU/AMU/FMU*	-.009	.052	-.166	.868	.991	.893	1.100
OU/FMU	.026	.050	.526	.601	1.027	.929	1.134
OU/AMU	-.031	.045	-.698	.487	.969	.886	1.060
OU only	0[b]	.	.	.	1.000	.	.
ZConsultantsFTE/birth	.028	.023	1.239	.218	1.028	.983	1.075
ZDoctors FTE/birth	-.034	.021	-1.571	.119	.967	.927	1.009
ZMidwives FTE/birth	-.042	.020	-2.080	.040	.959	.922	.998
ZHCA FTE/birth	-.004	.018	-.233	.816	.996	.961	1.032

b: Reference categories
OU (Obstetric Unit); AMU (Alongside Midwifery Unit); FMU (Freestanding Midwifery Unit)

Table 6.1-3: Random Effect (Level-1 & Level-2 predictors Emergency CS with missing data model)

					95% Confidence Interval	
Random Effect Covariance	Estimate	Std. Error	Z	Sig.	Lower	Upper
Var (Intercept)	.044	.006	7.306	<.001	.034	.058

Covariance Structure: Variance components
Subject Specification: Trust ID

6.1.4.2 TABLES FOR INSTRUMENTAL DELIVERY MISSING DATA ANALYSIS

Table 6.1-4: Fixed Effects Level-1 & Level-2 predictors, Instrumental Delivery, missing data model

N of cases 261,468; N of trusts 143

Source	F	Nom df1	Denom df2	Sig.
Corrected Model	77.172	31	439	<.001
Maternal Age	176.139	6	261436	<.001
Ethnicity	69.964	4	261436	<.001
IMD	13.465	5	261436	<.001
NICE Risk	27.069	1	261436	<.001
Gestational Age	92.136	2	185518	<.001
Birth Weight	42.959	3	261436	<.001
University Hospital	1.334	1	162	.250
London Trust	.146	1	74	.703
Foundation Trust , Sept 2010	1.811	1	93	.182
Trust OU/AMU/FMU	.765	3	118	.516
ZConsultants FTE/birth	6.385	1	95	.013
ZDoctors FTE/birth	.182	1	105	.670
ZMidwives FTE/birth	1.538	1	138	.217
ZHCA FTE/birth	.043	1	104	.836

Table 6.1-5: Fixed Coefficients, Level-1 & Level-2 predictors, Instrumental Delivery, missing data model

N of cases 261,468; N of trusts 143

						95% Confidence Interval for Odds Ratio	
Model Term	Coefficient	Std. Error	t	Sig.	Odds Ratio	Lower	Upper
Intercept	-1.623	.037	-43.815	<.001	.197	.183	.212
Level-1 predictors							
Maternal Age							
Missing	-.157	.154	-1.024	.306	.854	.632	1.155
40-44	.319	.036	8.758	<.001	1.376	1.281	1.478
35-39	.528	.027	19.860	<.001	1.695	1.609	1.786
30-34	.582	.020	29.140	<.001	1.789	1.720	1.860

25-29	.485	.019	25.225	<.001	1.623	1.563	1.686
20-24	.250	.019	13.159	<.001	1.284	1.237	1.333
15-19	0[b]	.	.	.	1.000	.	.
Ethnicity							
Missing	-.064	.022	-2.895	.004	.938	.898	.979
Mixed/Chinese/Other	-.131	.024	-5.521	<.001	.877	.837	.919
Black-Caribbean	-.753	.053	-14.097	<.001	.471	.424	.523
Asian	.051	.020	2.581	.010	1.053	1.012	1.095
White	0[b]	.	.	.	1.000	.	.
IMD in quintiles							
Missing	-.115	.075	-1.533	.125	.892	.770	1.033
5 Most Deprived	-.148	.019	-7.746	<.001	.862	.831	.895
4	-.109	.019	-5.822	<.001	.897	.865	.930
3	-.046	.016	-2.834	.005	.955	.926	.986
2	-.041	.015	-2.638	.008	.960	.932	.990
1 Least Deprived	0[b]	.	.	.	1.000	.	.
NICE Risk							
High Risk	-.085	.016	-5.203	<.001	.919	.890	.948
Low Risk	0[b]	.	.	.	1.000	.	.
Gestational Age							
Missing	-.041	.050	-.822	.411	.959	.869	1.059
>=42 weeks	.276	.021	13.406	<.001	1.318	1.265	1.372
37-41 weeks	0[b]	.	.	.	1.000	.	.
Birth Weight							
Missing	-.098	.069	-1.421	.155	.906	.791	1.038
>4500g	-.109	.045	-2.453	.014	.896	.821	.978
<2500g	-.408	.036	-11.211	<.001	.665	.620	.714
2500g-4500g	0[b]	.	.	.	1.000	.	.
Level-2 predictors							
University Hospital	.058	.050	1.155	.250	1.060	.960	1.170
Not Uni-Hospital	0[b]	.	.	.	1.000	.	.
London Trust	-.020	.054	-.383	.703	.980	.881	1.090
Not London Trust	0[b]	.	.	.	1.000	.	.
Foundation Trust	.052	.039	1.346	.182	1.053	.976	1.137
Not Foundation Trust	0[b]	.	.	.	1.000	.	.
OU/AMU/FMU*	.024	.060	.403	.688	1.024	.910	1.153
OU/FMU	-.053	.062	-.854	.395	.948	.838	1.073
OU/AMU	.051	.049	1.027	.306	1.052	.954	1.159
OU only	0[b]	.	.	.	1.000	.	.

ZConsultantsFTE/birth	.058	.023	2.527	.013	1.060	1.013	1.109
ZDoctors FTE/birth	.009	.022	.427	.670	1.009	.966	1.055
ZMidwives FTE/birth	-.029	.024	-1.240	.217	.971	.926	1.018
ZHCA FTE/birth	.004	.021	.207	.836	1.004	.964	1.046

b: Reference categories
OU (Obstetric Unit); AMU (Alongside Midwifery Unit); FMU (Freestanding Midwifery Unit)

Table 6.1-6: Random Effect (Level-1 & Level-2 predictors Instrumental Delivery with missing data model)

					95% Confidence Interval	
Random Effect Covariance	Estimate	Std. Error	Z	Sig.	Lower	Upper
Var(Intercept)	.052	.007	7.496	<.001	.040	.068

Covariance Structure: Variance components
Subject Specification: Trust ID

6.1.4.3 TABLES FOR NORMAL BIRTH MISSING DATA ANALYSIS

Table 6.1-7: Fixed Effects Level-1 & Level-2 predictors, Normal Birth, missing data model

N of cases 214,920; N of trusts 127

Source	F	Nom df1	Denom df2	Sig.
Corrected Model	187.484	31	249	<.001
Maternal Age	135.998	6	214888	<.001
Ethnicity	49.521	4	214888	<.001
IMD	7.008	5	214888	<.001
NICE Risk	2750.851	1	214888	<.001
Gestational Age	107.981	2	214888	<.001
Birth Weight	96.130	3	214888	<.001
University Hospital	2.603	1	140	.109
London Trust	1.764	1	91	.187
Foundation Trust , Sept 2010	1.017	1	79	.316
Trust OU/AMU/FMU	2.513	3	86	.064
ZConsultants FTE/birth	.280	1	59	.599
ZDoctors FTE/birth	1.683	1	96	.198
ZMidwives FTE/birth	5.126	1	63	.027
ZHCA FTE/birth	.146	1	34	.705

Table 6.1-8: Fixed Coefficients, Level-1 & Level-2 predictors, Normal Birth, missing data model

N of cases 214,920; N of trusts 127

						95% Confidence Interval for Odds Ratio	
Model Term	Coefficient	Std. Error	t	Sig.	Odds Ratio	Lower	Upper
Intercept	-1.431	.055	-26.008	<.001	.239	.215	.266
Level-1 predictors							
Maternal Age							
Missing	-.209	.203	-1.031	.303	.811	.545	1.208
40-44	-.954	.067	-14.326	<.001	.385	.338	.439
35-39	-.820	.038	-21.469	<.001	.440	.409	.475
30-34	-.685	.026	-26.547	<.001	.504	.479	.530

25-29	-.477	.021	-22.244	<.001	.621	.595	.647
20-24	-.224	.021	-10.936	<.001	.799	.768	.832
15-19	0[b]	.	.	.	1.000	.	.
Ethnicity							
Missing	.108	.023	4.803	<.001	1.114	1.066	1.165
Mixed/Chinese/Other	.043	.030	1.409	.159	1.044	.983	1.107
Black-Caribbean	.228	.037	6.172	<.001	1.256	1.168	1.351
Asian	-.195	.023	-8.451	<.001	.823	.786	.861
White	0[b]	.	.	.	1.000	.	.
IMD in quintiles							
Missing	.076	.103	.736	.462	1.079	.882	1.319
5 Most Deprived	.132	.024	5.581	<.001	1.141	1.089	1.195
4	.099	.022	4.525	<.001	1.104	1.057	1.152
3	.042	.019	2.255	.024	1.043	1.006	1.082
2	.043	.018	2.388	.017	1.044	1.008	1.082
1 Least Deprived	0[b]	.	.	.	1.000	.	.
NICE Risk							
Low Risk	1.548	.030	52.449	<.001	4.701	4.437	4.981
High Risk	0[b]	.	.	.	1.000	.	.
Gestational Age							
Missing	.127	.063	2.023	.043	1.135	1.004	1.284
>=42 weeks	-.871	.061	-14.230	<.001	.419	.371	.472
37-41 weeks	0[b]	.	.	.	1.000	.	.
Birth Weight							
Missing	.084	.186	.452	.652	1.088	.755	1.566
>4500g	-1.027	.065	-15.911	<.001	.358	.315	.406
<2500g	.201	.036	5.532	<.001	1.223	1.139	1.313
2500g-4500g	0[b]	.	.	.	1.000	.	.
Level-2 predictors							
University Hospital	-.087	.054	-1.613	.109	.917	.825	1.020
Not Uni-Hospital	0[b]	.	.	.	1.000	.	.
London Trust	-.087	.065	-1.328	.187	.917	.805	1.044
Not London Trust	0[b]	.	.	.	1.000	.	.
Foundation Trust	-.042	.042	-1.008	.316	.959	.882	1.042
Not Foundation Trust	0[b]	.	.	.	1.000	.	.
OU/AMU/FMU*	.106	.068	1.564	.121	1.112	.972	1.272
OU/FMU	.160	.064	2.515	.014	1.174	1.034	1.332
OU/AMU	.032	.053	.601	.549	1.032	.929	1.147
OU only	0[b]	.	.	.	1.000	.	.

ZConsultantsFTE/birth	-.012	.023	-.529	.599	.988	.944	1.034
ZDoctors FTE/birth	-.032	.025	-1.297	.198	.969	.923	1.017
ZMidwives FTE/birth	.049	.022	2.264	.027	1.051	1.006	1.097
ZHCA FTE/birth	-.007	.018	-.382	.705	.993	.958	1.030

b: Reference categories
OU (Obstetric Unit); AMU (Alongside Midwifery Unit); FMU (Freestanding Midwifery Unit)

Table 6.1-9: Random Effect (Level-1 & Level-2 predictors) Normal Birth with missing data model)

					95% Confidence Interval	
Random Effect Covariance	Estimate	Std. Error	Z	Sig.	Lower	Upper
Var (Intercept)	.056	.008	7.071	<.001	.042	.074

Covariance Structure: Variance components
Subject Specification: Trust ID

6.2 CROSS-LEVEL INTERACTIONS

A second sensitivity analysis was performed to try and answer:

Q4: Are the effects of standardized staff FTE/birth ratios on the three outcomes different for low-risk and high-risk women?

Cross-level interaction effects between each standardized staff FTE/birth ratio and individual-level NICE clinical risk were explored. A significant interaction coefficient would indicate that the effects of staffing on the outcomes vary with women's clinical risk.

The cross-level interaction models were performed on subsamples of 111 (out of 143) trusts for emergency CS and instrumental delivery and 104 (out of 127) trusts for normal birth, which resulted from exclusion of trusts with 100% missing data and trusts with more than 20% missing data on the following variables:

- gestational age (10 trusts – 100% missing and 16 trusts >20% missing data)
- infant birth weight (7 trusts – 100% missing and 11 trusts >20% missing data)
- ethnicity (2 trusts >20% missing data)
- IMD (1 trust with >20% missing data).

Despite selecting trusts with more complete data on the main predictors, the analysed samples in the interaction models for emergency CS and instrumental delivery were reduced by 10.3%; and by 8.4% in the interaction model for normal birth.

The Tables with results from the cross-level interaction models for the three outcomes, based on the above selection are presented at the end of this chapter.

Before testing the interactions to the above selection of trusts, interaction models were also run on the initial 143 trusts for emergency CS and instrumental delivery

and the 127 trusts for normal birth. The models were unable to calculate F-statistics for fixed effects for several of the standardized staff FTE/birth ratios and their interaction effects with NICE Risk. For the cross-level interaction models presented below, only normal birth model was unable to calculate F-statistics for fixed effects of the standardized HCA FTE/birth ratio and its interaction with NICE Risk.

6.2.1 EMERGENCY CS – CROSS-LEVEL INTERACTION ANALYSIS

The results from the cross-level interaction model for emergency CS are presented in Table 6.2-1, Table 6.2-2 and Table 6.2-3 at the end of this section. The fixed effects and fixed coefficients were very similar to the main results for emergency CS presented in Chapter 5. The only difference was the slightly stronger effect of doctors on emergency CS, as well as the significance of the doctors/clinical risk interaction. The standardized doctors FTE/birth ratio was negatively related to the probability of emergency CS (OR=0.930, $p<.05$), while high risk women were positively related to emergency CS (OR=2.317, $p<.001$) compared to low risk women Table 6.2-2. The interaction effect suggested that if a woman classified as NICE high risk was in a trust with 1 SD above the grand mean for doctors FTE/birth ratio, the odds of having an emergency CS were multiplied by 1.051 (CI 1.010-1.095) or increased by 5.1% (Table 6.2-2), which indicated higher odds of having an emergency CS for a woman who was NICE high risk compared to similar woman in a trust with average ratio of doctors FTE/birth, adjusted for other predictors in the model. The variance of the cross-level interaction model for emergency CS was 0.047 (SE 0.007) Table 6.2-3.

6.2.2 INSTRUMENTAL DELIVERY – CROSS-LEVEL INTERACTION ANALYSIS

The results from the cross-level interaction model for instrumental delivery are presented in Table 6.2-4, Table 6.2-5 and Table 6.2-6 at the end of this section. Again the fixed effects and fixed coefficients were similar to the main results presented for instrumental delivery earlier in Chapter 5. None of the interaction effects were significant in the model for instrumental delivery. The variance of the cross-level interaction model for instrumental delivery was 0.049 (SE 0.008), Table 6.2-6.

6.2.3 NORMAL BIRTH – CROSS-LEVEL INTERACTION ANALYSIS

The results from the cross-level interaction model for normal birth are presented in Table 6.2-7, Table 6.2-8 and Table 6.2-9. The fixed effects and fixed coefficients were similar to the main results presented for normal birth in Chapter 5. There was an indication that giving birth in a London trust reduced the odds of normal birth (OR=0.869, CI 0.761-0.993, p<.05), Table 6.2-8. The positive and significant interaction effect between NICE risk and standardized midwives FTE/birth ratio, suggested that if a woman classified NICE low risk in a trust with 1 SD above the grand mean for midwives FTE/birth ratio, the odds of having a normal birth were multiplied by 1.076 (CI 1.017-1.138) or increased by 7.6% (Table 6.2-8), which indicates higher odds of having a normal birth for a woman who was NICE low risk compared to similar woman who was in a trust with more average ratio of midwives FTE/birth, adjusted for other predictors in the model. The variance of the cross-level interaction model for normal birth CS was 0.053 (SE 0.008), Table 6.2-9.

6.2.4 CONCLUSIONS

The cross-level interaction effects between the staffing variables and women's clinical risk indicated: a 5.1% increase in the odds of an emergency CS for women who were high risk (HR) in trusts with 1 SD above the grand mean of FTE doctors/birth ratio (OR=1.051, CI 1.010-1.095, p<.05), compared to similar HR

women in trusts with average ratio of doctors FTE/birth; and a 7.6% increase in the odds of a normal birth for low risk (LR) women in trusts with 1 SD above the grand mean of FTE midwives/birth ratio (OR = 1.076, CI 1.017-1.138, p<0.05), compared to similar LR women in trusts with average ratio of midwives FTE/birth, adjusted for the other predictors in the models. In the cross-level interaction models, for all women irrespective of risk, the standardized doctors FTE/birth ratio was negatively related to emergency CS (OR=0.930, CI 0.878-0.986, p<.05); the standardized consultants FTE/birth ratio was positively related to instrumental delivery (OR=1.076, CI 1.021-1.135, p<.05); the standardized midwives FTE/birth ratio was negatively related to instrumental delivery (OR=0.936, CI 0.886-0.988, p<.05) and the standardized doctors FTE/birth ratio was negatively related to normal birth (OR=0.930, CI 0.868-0.996, p<.05), while giving birth in a London trust reduced the odds of normal birth (OR=0.869, CI 0.761-0.993, p<.05).

6.2.5 TABLES FROM CROSS-LEVEL INTERACTION ANALYSES

6.2.5.1 TABLES FOR CROSS-LEVEL INTERACTIONS IN EMERGENCY CS MODEL

Table 6.2-1: Fixed Effects, Cross-Level Interactions in Emergency CS Model

N of cases 177,560; N of trusts 111

Source	F	Nom df1	Denom df2	Sig.
Corrected Model	235.901	30	224	<.001
Maternal Age	272.477	5	177529	<.001
Ethnicity	103.534	3	177529	<.001
IMD	4.910	4	177529	.001
NICE Risk	2031.421	1	177529	<.001
Gestational Age	239.136	1	177529	<.001
Birth Weight	284.785	2	177529	<.001
University Hospital	.066	1	76	.798
London Trust	1.124	1	88	.292
Foundation Trust , Sept 2010	.019	1	74	.890
Trust OU/AMU/FMU	.397	3	58	.756
ZConsultants FTE/birth	.494	1	90	.484
ZDoctors FTE/birth	3.835	1	78	.054
ZMidwives FTE/birth	1.057	1	46	.309
ZHCA FTE/birth	.198	1	56	.658
NICE Risk * ZConsultants FTE/birth	1.978	1	177529	.160
NICE Risk * ZDoctors FTE/birth	5.857	1	177529	.016
NICE Risk * ZMidwives FTE/birth	.645	1	177529	.422
NICE Risk * ZHCA FTE/birth	.034	1	177529	.853

Table 6.2-2: Fixed Coefficients, Cross-Level Interactions in Emergency CS Model

N of cases 177,560 cases; N of trusts 111

Model Term	Coefficient	Std. Error	t	Sig.	Odds Ratio	95% Confidence Interval for Odds Ratio Lower	Upper
Intercept	-2.601	.053	-49.079	<.001	.074	.067	.082
Level-1 predictors							
Maternal Age							
40-44	1.073	.051	21.058	<.001	2.925	2.647	3.233
35-39	1.000	.038	26.432	<.001	2.719	2.525	2.929
30-34	.770	.032	23.999	<.001	2.159	2.027	2.299
25-29	.569	.033	17.060	<.001	1.766	1.654	1.885
20-24	.309	.033	9.349	<.001	1.362	1.277	1.453
15-19	0[b]	.	.	.	1.000	.	.
Ethnicity							
Mixed/Chinese/Other	.147	.027	5.403	<.001	1.159	1.098	1.222
Black-Caribbean	.503	.030	16.985	<.001	1.653	1.560	1.752
Asian	.231	.028	8.242	<.001	1.260	1.193	1.332
White	0[b]	.	.	.	1.000	.	.
IMD in quintiles							
5 Most Deprived	.100	.024	4.156	<.001	1.105	1.054	1.159
4	.072	.022	3.297	.001	1.075	1.030	1.122
3	.063	.019	3.298	.001	1.065	1.026	1.106
2	.047	.020	2.294	.022	1.048	1.007	1.091
1 Least Deprived	0[b]	.	.	.	1.000	.	.
NICE Risk							
NICE High Risk	.840	.019	45.071	<.001	2.317	2.234	2.403
NICE Low Risk	0[b]	.	.	.	1.000	.	.
Gestational Age							
>=42 weeks	.490	.032	15.464	<.001	1.632	1.534	1.737
37-41 weeks	0[b]	.	.	.	1.000	.	.
Birth Weight							
>4500g	1.167	.051	22.882	<.001	3.212	2.907	3.550
<2500g	.190	.036	5.314	<.001	1.209	1.128	1.297
2500g-4500g	0[b]	.	.	.	1.000	.	.
Level-2 predictors							
University Hospital	-.013	.049	-.257	.798	.988	.896	1.088

Not Uni-Hospital	0[b]	.	.	.	1.000	.	.
London Trust	.070	.067	1.060	.292	1.073	.940	1.225
Not London Trust	0[b]	.	.	.	1.000	.	.
Foundation Trust	-.006	.043	-.139	.890	.994	.912	1.083
Not Foundation Trust	0[b]	.	.	.	1.000	.	.
OU/AMU/FMU*	-.012	.066	-.180	.858	.988	.866	1.128
OU/FMU	.051	.053	.963	.342	1.053	.945	1.173
OU/AMU	.010	.057	.173	.863	1.010	.902	1.130
OU only	0[b]	.	.	.	1.000	.	.
ZConsultantsFTE/birth	.034	.030	1.130	.261	1.034	.975	1.097
ZDoctors FTE/birth	-.072	.029	-2.473	.014	.930	.878	.986
ZMidwives FTE/birth	-.033	.026	-1.295	.199	.967	.919	1.018
ZHCA FTE/birth	-.007	.023	-.314	.754	.993	.948	1.040
Cross-Level Interaction Terms							
NICE HR*ZConsultantsFTE/birth	-.029	.020	-1.406	.160	.972	.934	1.011
NICE LR*ZConsultantsFTE/birth	0[b]	.	.	.	1.000	.	.
NICE HR*ZDoctors FTE/birth	.050	.021	2.420	.016	1.051	1.010	1.095
NICE LR*ZDoctors FTE/birth	0[b]	.	.	.	1.000	.	.
NICE HR*ZMidwives FTE/birth	.020	.024	.803	.422	1.020	.972	1.070
NICE LR*ZMidwives FTE/birth	0[b]	.	.	.	1.000	.	.
NICE HR*ZHCA FTE/birth	-.003	.017	-.185	.853	.997	.963	1.031
NICE LR*ZHCA FTE/birth	0[b]	.	.	.	1.000	.	.

b: Reference categories
OU (Obstetric Unit); AMU (Alongside Midwifery Unit); FMU (Freestanding Midwifery Unit)
NICE HR (NICE High Risk), NICE LR (NICE Low Risk)

Table 6.2-3: Random Effect, Cross-Level Interactions in Emergency CS Model

Random Effect

					95% Confidence Interval	
Random Effect Covariance	Estimate	Std. Error	Z	Sig.	Lower	Upper
Var (Intercept)	.047	.007	6.345	<.001	.035	.064

Covariance Structure: Variance components
Subject Specification: Trust ID

6.2.5.2 TABLES FOR CROSS-LEVEL INTERACTION IN INSTRUMENTAL DELIVERY MODEL

Table 6.2-4: Fixed Effects, Cross-Level Interactions in Instrumental Delivery Model

N of cases 177,560; N of trusts 111

Source	F	Nom df1	Denom df2	Sig.
Corrected Model	76.601	30	277	<.001
Maternal Age	146.854	5	177529	<.001
Ethnicity	125.811	3	177529	<.001
IMD	7.882	4	177529	<.001
NICE Risk	22.320	1	177529	<.001
Gestational Age	191.520	1	177529	<.001
Birth Weight	46.737	2	177529	<.001
University Hospital	.223	1	141	.638
London Trust	.003	1	65	.956
Foundation Trust , Sept 2010	1.346	1	64	.250
Trust OU/AMU/FMU	.198	3	97	.897
ZConsultants FTE/birth	6.668	1	63	.012
ZDoctors FTE/birth	.566	1	64	.455
ZMidwives FTE/birth	4.901	1	84	.030
ZHCA FTE/birth	.549	1	45	.463
NICE Risk * ZConsultants FTE/birth	.813	1	177529	.367
NICE Risk * ZDoctors FTE/birth	2.020	1	177529	.155
NICE Risk * ZMidwives FTE/birth	.743	1	177529	.389
NICE Risk * ZHCA FTE/birth	1.241	1	177529	.265

Table 6.2-5: Fixed Coefficients, Cross-Level Interactions in Instrumental Delivery Model

N of cases 177,560; N of trusts 111

						95% Confidence Interval for Odds Ratio	
Model Term	Coefficient	Std. Error	t	Sig.	Odds Ratio	Lower	Upper
Intercept	-1.658	.042	-39.129	<.001	.190	.175	.207
Level-1 predictors							

Maternal Age							
40-44	.361	.043	8.346	<.001	1.435	1.318	1.562
35-39	.561	.031	17.957	<.001	1.753	1.649	1.864
30-34	.607	.025	23.991	<.001	1.834	1.745	1.927
25-29	.509	.022	22.920	<.001	1.664	1.593	1.738
20-24	.279	.023	11.968	<.001	1.322	1.263	1.384
15-19	0[b]	.	.	.	1.000	.	.
Ethnicity							
Mixed/Chinese/Other	-.137	.028	-4.864	<.001	.872	.825	.921
Black-Caribbean	-.763	.044	-17.392	<.001	.466	.428	.508
Asian	.057	.022	2.572	<.010	1.058	1.014	1.105
White	0[b]	.	.	.	1.000	.	.
IMD in quintiles							
5 Most Deprived	-.126	.023	-5.417	<.001	.881	.842	.922
4	-.084	.022	-3.880	<.001	.919	.881	.959
3	-.035	.019	-1.859	.063	.966	.931	1.002
2	-.032	.018	-1.759	.079	.968	.934	1.004
1 Least Deprived	0[b]	.	.	.	1.000	.	.
NICE Risk							
NICE High Risk	-.088	.019	-4.724	<.001	.916	.883	.950
NICE Low Risk	0[b]	.	.	.	1.000	.	.
Gestational Age							
>=42 weeks	.293	.021	13.839	<.001	1.341	1.286	1.397
37-41 weeks	0[b]	.	.	.	1.000	.	.
Birth Weight							
>4500g	-.116	.050	-2.345	.019	.890	.808	.981
<2500g	-.389	.042	-9.288	<.001	.678	.624	.736
2500g-4500g	0[b]	.	.	.	1.000	.	.
Level-2 predictors							
University Hospital	.027	.058	.472	.638	1.028	.917	1.152
Not Uni-Hospital	0[b]	.	.	.	1.000	.	.
London Trust	.003	.063	.056	.956	1.003	.885	1.137
Not London Trust	0[b]	.	.	.	1.000	.	.
Foundation Trust	.049	.042	1.160	.250	1.050	.965	1.143
Not Foundation Trust	0[b]	.	.	.	1.000	.	.
OU/AMU/FMU*	.050	.070	.713	.478	1.051	.914	1.209
OU/FMU	-.003	.067	-.047	.963	.997	.872	1.139
OU/AMU	.028	.061	.452	.652	1.028	.911	1.161
OU only	0[b]	.	.	.	1.000	.	.

ZConsultantsFTE/birth	.074	.027	2.769	.007	1.076	1.021	1.135
ZDoctors FTE/birth	.001	.025	.043	.966	1.001	.953	1.051
ZMidwives FTE/birth	-.067	.028	-2.420	.017	.936	.886	.988
ZHCA FTE/birth	.007	.021	.323	.748	1.007	.966	1.049
Cross-Level Interaction Terms							
NICE HR*ZConsultantsFTE/birth	-.016	.017	-.902	.367	.985	.952	1.018
NICE LR*ZConsultantsFTE/birth	0[b]	.	.	.	1.000	.	.
NICE HR*ZDoctors FTE/birth	.033	.023	1.421	.155	1.033	.988	1.081
NICE LR*ZDoctors FTE/birth	0[b]	.	.	.	1.000	.	.
NICE HR*ZMidwives FTE/birth	.014	.017	.862	.389	1.014	.982	1.048
NICE LR * ZMidwives FTE/birth	0[b]	.	.	.	1.000	.	.
NICE HR * ZHCA FTE/birth	.015	.014	1.114	.265	1.015	.989	1.043
NICE LR * ZHCA FTE/birth	0[b]	.	.	.	1.000	.	.

b: Reference categories
*OU (Obstetric Unit); AMU (Alongside Midwifery Unit); FMU (Freestanding Midwifery Unit)
NICE HR (NICE High Risk), NICE LR (NICE Low Risk)

Table 6.2-6: Random Effect, Cross-Level Interaction in Instrumental Delivery Model

Random Effect

					95% Confidence Interval	
Random Effect Covariance	Estimate	Std. Error	Z	Sig.	Lower	Upper
Var(Intercept)	.049	.008	6.417	<.001	.036	.067

Covariance Structure: Variance components
Subject Specification: Trust ID

6.2.5.3 TABLES FOR CROSS-LEVEL INTERACTIONS IN NORMAL BIRTH MODEL

Table 6.2-7: Fixed Effects, Cross-Level Interactions in Normal Birth Model

N of cases 162,337; N of trusts 104

Source	F	Nom df1	Denom df2	Sig.
Corrected Model	179.371	29	143	<.001
Maternal Age	140.859	5	162307	<.001
Ethnicity	34.987	3	162307	<.001
IMD	5.184	4	162307	<.001
NICE Risk	2183.748	1	162307	<.001
Gestational Age	157.944	1	162307	<.001
Birth Weight	135.457	2	162307	<.001
University Hospital	1.455	1	103	.231
London Trust	4.491	1	49	.039
Foundation Trust , Sept 2010	.420	1	64	.519
Trust OU/AMU/FMU	1.302	3	69	.281
ZConsultants FTE/birth	.192	1	49	.663
ZDoctors FTE/birth	3.385	1	80	.070
ZMidwives FTE/birth	2.253	1	48	.140
ZHCA FTE/birth	.211	1	34	.649
NICE Risk * ZConsultants FTE/birth	.187	1	162307	.665
NICE Risk * ZDoctors FTE/birth	2.379	1	162307	.123
NICE Risk * ZMidwives FTE/birth	6.541	1	162307	.011

**The Model was unable to calculate F statistics for the interaction of standardized HCA FTE/birth ratio and NICE Risk

Table 6.2-8: Fixed Coefficients, Cross-Level Interactions in Normal Birth Model

N of cases 162,337; N of trusts 104

						95% Confidence Interval for Odds Ratio	
Model Term	Coeff.	Std. Error	t	Sig.	Odds Ratio	Lower	Upper
Intercept	-1.412	.058	-24.200	<.001	.244	.217	.273
Level-1 predictors							
Maternal Age							
40-44	-.966	.077	-12.485	<.001	.381	.327	.443
35-39	-.845	.042	-20.086	<.001	.430	.396	.467
30-34	-.695	.029	-23.891	<.001	.499	.471	.528

25-29	-.491	.023	-20.939	<.001	.612	.585	.641
20-24	-.234	.023	-10.207	<.001	.791	.756	.828
15-19	0[b]	.	.	.	1.000	.	.
Ethnicity							
Mixed/Chinese/Other	.033	.031	1.079	<.281	1.034	.973	1.097
Black-Caribbean	.223	.039	5.676	<.001	1.250	1.157	1.351
Asian	-.180	.024	-7.451	<.001	.835	.796	.876
White	0[b]	.	.	.	1.000	.	.
IMD in quintiles							
5 Most Deprived	.110	.026	4.184	<.001	1.117	1.060	1.176
4	.091	.026	3.475	.001	1.095	1.040	1.153
3	.031	.022	1.453	.146	1.032	.989	1.076
2	.042	.021	2.040	.041	1.043	1.002	1.086
1 Least Deprived	0[b]	.	.	.	1.000	.	.
NICE Risk							
NICE Low Risk	1.557	.033	46.731	<.001	4.746	4.446	5.067
NICE High Risk	0[b]	.	.	.	1.000	.	.
Gestational Age							
>=42 weeks	-.846	.067	-12.568	<.001	.429	.376	.490
37-41 weeks	0[b]	.	.	.	1.000	.	.
Birth Weight							
>4500g	-1.057	.066	-15.988	<.001	.347	.305	.395
<2500g	.169	.039	4.317	<.001	1.184	1.096	1.278
2500g-4500g	0[b]	.	.	.	1.000	.	.
Level-2 predictors							
University Hospital	-.070	.058	-1.206	.231	.933	.831	1.046
Not Uni-Hospital	0[b]	.	.	.	1.000	.	.
London Trust	-.140	.066	-2.119	.039	.869	.761	.993
Not London Trust	0[b]	.	.	.	1.000	.	.
Foundation Trust	-.029	.046	-.648	.519	.971	.887	1.063
Not Foundation Trust	0[b]	.	.	.	1.000	.	.
OU/AMU/FMU*	.113	.076	1.473	.145	1.119	.961	1.303
OU/FMU	.100	.065	1.537	.130	1.105	.970	1.259
OU/AMU	.024	.063	.383	.703	1.024	.904	1.161
OU only	0[b]	.	.	.	1.000	.	.
ZConsultantsFTE/birth	.018	.034	.529	.598	1.018	.952	1.089
ZDoctors FTE/birth	-.073	.035	-2.081	.039	.930	.868	.996
ZMidwives FTE/birth	.001	.033	.034	.973	1.001	.937	1.069
ZHCA FTE/birth	-.009	.019	-.459	.649	.991	.953	1.031

Cross-Level Interaction Terms							
NICE LR*ZConsultantsFTE/birth	-.014	.032	-.433	.665	.986	.926	1.050
NICE HR*ZConsultantsFTE/birth	0[b]	.	.	.	1.000	.	.
NICE LR*ZDoctors FTE/birth	.048	.031	1.543	.123	1.049	.987	1.115
NICE HR*ZDoctors FTE/birth	0[b]	.	.	.	1.000	.	.
NICE LR*ZMidwives FTE/birth	.073	.029	2.558	.011	1.076	1.017	1.138
NICE HR*ZMidwives FTE/birth	0[b]	.	.	.	1.000	.	.

b: Reference categories
*OU (Obstetric Unit); AMU (Alongside Midwifery Unit); FMU (Freestanding Midwifery Unit)
**NICE HR (NICE High Risk), NICE LR (NICE Low Risk)
***The Model was unable to calculate F statistics for the interaction of standardized HCA FTE/birth ratio and NICE Risk

Table 6.2-9: Random Effect, Cross-Level Interaction in Normal Birth Model

Random Effect						
					95% Confidence Interval	
Random Effect Covariance	Estimate	Std. Error	Z	Sig.	Lower	Upper
Var (Intercept)	.053	.008	6.311	<.001	.039	.072

Covariance Structure: Variance components
Subject Specification: Trust ID

7 CHAPTER 7 DISCUSSION

7.1 INTRODUCTION

This study aimed to investigate the relationship between maternity staffing levels, certain structural characteristics of NHS trusts and three selected birth modes – two process measures (emergency CS and instrumental delivery) and one outcome measure (normal birth, which was a derived indicator with a specific definition) in NHS trusts with maternity services in England in 2010/11.

The study aimed to examine the hypothesis that trusts and workforce differences in the provision of maternity care contribute to variations in the three indicators across trusts. Specifically it aimed to answer the following questions:

1. What is the relationship between emergency CS, instrumental delivery and normal birth and maternity staffing levels (FTE/birth ratios) in NHS trusts in England in 2010/11, after accounting for maternal socio-demographic characteristics, individual clinical risk and structural characteristics including type and configuration of trusts?
2. Is this relationship, if any, different across the three indicators?
3. To what extent do maternal socio-demographic characteristics and clinical risk factors determine the likelihood of emergency CS, instrumental delivery and normal birth?
4. To what extent are variations in emergency CS, instrumental delivery and normal birth between NHS trusts with maternity services in England explained by differences in staffing levels and trusts' characteristics?

The analyses have yielded some important findings, which are discussed below. The policy relevance of findings are also considered.

Certain limitations regarding the use of routinely collected data and the level at which staffing data were available are also presented.

This Chapter discusses the study results (presented in Chapter 5 and Chapter 6) and relates these to the research aims and hypothesis and previous literature on staffing, emergency CS, instrumental delivery and normal birth.

7.2 PRINCIPAL FINDINGS

This study found an independent effect of maternity staffing on the three selected indicators, based on a sample of women aged 15-44, who were nulliparous and had a term (>=37 weeks), singleton, live birth in England, after controlling for maternal socio-demographic characteristics and individual clinical risk.

Specifically, for all women irrespective of risk, more midwives increased the chance of normal birth. More consultants increased the likelihood of instrumental delivery for all women.

The analyses of the multiplicative effect of staffing and clinical risk indicated higher odds of having an emergency CS for women who met NICE (2007) high risk criteria in a trust with more doctors compared to similar women in a trust with average doctors/birth ratios; and higher odds of having a normal birth for women who met NICE (2007) low risk criteria in a trust with more midwives compared to similar women in a trust with average midwives/birth ratios.

The study did not find any statistically significant relationship between healthcare assistants and any of the indicators.

There was some indication that women giving birth in trusts which had both obstetric and freestanding midwifery units may have improved chances of normal birth, while women giving birth in a London trust had reduced odds of normal birth. Overall there were almost no associations between the selected trusts characteristics and the three indicators.

More than anything else, women's outcomes were determined by their characteristics and clinical risk. Overall age, clinical risk, birth weight and gestational age had a

strong and significant relationship with all outcomes. Ethnicity and IMD, though significant were weaker predictors of the three indicators.

Despite some significant independent effects of staffing on the three indicators, this study found that only a small percentage (1%-2%) of the total variability in emergency CS, instrumental delivery and normal birth was attributable to variations between trusts. Adjusting for socio-demographics, clinical risk, staffing and trusts characteristics did not reduce substantially the unexplained trust-level variations; indeed in some cases their inclusion increased the between-trust variances.

7.3 DISCUSSION

7.3.1 SELECTION OF INDICATORS

The three indicators were selected because of their policy relevance; availability in HES; reliability in coding[106]; and to compare results across each of them as well as with other studies.

These indicators also represented various dimensions of care. For example, normal birth is considered a desired outcome for majority of healthy women and is more closely related to midwifery care for low-risk women; while the more interventionist approaches to labour (emergency CS and instrumental delivery) may have positive or negative consequences for women and their infants and result from care directly provided by medical staff, or may be related indirectly to care provided by midwives. The more interventionist approaches were also relevant because of their link to maternal morbidities, neonatal trauma and complications; length of hospital stay; rising rates (emergency CS); rising costs; litigations; woman/infant wellbeing; and impact on women's decisions for and outcomes of future pregnancies. Instrumental delivery and emergency CS though investigated independently in the current study

[106] Mode of delivery was well recorded in HES, with good internal agreement between HES OPCS delivery codes and maternity tail (Knight et al. 2013; RCOG 2013).

are not always independent of each other events, for example a failed instrumental delivery or unwillingness to perform instrumental delivery may both contribute to an increase and variation in emergency CS rates. These two process measures could also be seen as outcomes of prior interventions, such as induction or epidural.

Women who have had a caesarean section may also experience poorer psychological wellbeing related to post-traumatic stress disorder and the early cessation of breast feeding (Olde et al. 2006; Chalmers et al. 2010 and Smith 2007 cited in Essex et al. 2013). Instrumental vaginal births increase the risk of morbidity from pelvic floor injury compared with unassisted vaginal birth (RCOG 2011a) and there seemed to be similar psychological effects to those reported for caesarean section (Essex et al. 2013). Instrumental deliveries and CS represent a substantial financial cost to the UK National Health Service (Petrou and Glazener 2002 cited in Essex et al. 2013).

7.3.2 SELECTION OF SAMPLE

This study focused on a sample of women aged 15-44, who were nulliparous and had a term (>=37 weeks), singleton, live birth. Compared to the total population of women in HES, the descriptive statistics from the sample revealed a higher proportion of teen mothers (age 15-19); a lower proportion of older mothers (age>35); a higher proportion classified as NICE (2007) low risk than high risk; and a smaller proportion living in the most deprived quintile (IMD). Despite these characteristics, more women in this sample had operative births (emergency CS (19.3%) and instrumental delivery (21.8%)) and a smaller proportion had a normal birth (24.7%), compared to the total population of women in HES. Lower intervention rates and mainly caesarean section were considered important for the outcomes of future pregnancies for these women as there was a documented evidence of higher risk of uterine rupture (Fitzpatrick et al. 2012a), peripartum hysterectomy (Knight et al. 2008) and morbidly adherent placenta (Fitzpatrick et al. 2012b), all of which could contribute to increased perinatal mortality and morbidity rates. Certain conditions (eclampsia) occurred more often among younger and older mothers and in primiparous women (Knight 2005).

The focus of the study evolved from several factors. These include: an initial interest in exploring the relationship between midwifery staffing levels and outcomes sensitive to midwifery care; the realisation that midwifery staffing could not be investigated in isolation from other maternity staffing groups; rising interest in primary CS rates and particularly emergency CS rates; and the importance of first births for women. It was of interest to test a hypothesis that trusts and workforce differences in the provision of maternity care may explain previously observed variations in emergency CS rates (RCOG 2013; Bragg et al. 2010; NAO 2013) and instrumental deliveries (RCOG 2013; NAO 2013) across NHS trusts in England. The sample therefore was selected to test this hypothesis for first time mothers; and to allow for comparison between the effects of staffing on the three selected indicators and with other studies.

The sample of nulliparous women who had a term (>=37 weeks), singleton, live birth removed some of the clinical risk associated with multiple births, pre-term births, previous obstetric complications or prior caesarean births, thus creating a slightly more homogeneous study population. Despite this, variations in overall CS rates and emergency CS rates in a similar group were observed in other studies.

For example, a prospective regional study by Main et al. (2006) of all births, in a stratum of nulliparous, term, singleton, vertex population of women, over a period of three years in 20 hospitals in California, showed that although this was only one stratum of the obstetric population, it was a large stratum; there were large variations in CS rate within the stratum; the stratum was *"greatly affected by provider practices"*[107]; and better understanding and reduction in primary CS rates had the benefit of reducing the repeat CS rates. This thesis also found this was a large stratum: nearly 40% of women in 2010/11 in England were nulliparous, aged 15-44 years old with term, singleton, live births. And other studies have found large

[107] In Main et al. (2006) these practices were induction and early labour admission (latent phase), which was associated with dystocia and failure to progress resulting in potentially avoidable CS births.

variations in adjusted emergency CS rates for the stratum of nulliparous, 15-45 years old women with singleton, term, vertex births who had a spontaneous onset of labour. RCOG (2013) found these rates varied between 4.9% and 18.9% across 152 hospitals.

Main et al. (2006) and other studies formed the basis of the sample selection in the recently published RCOG Clinical Indicators Project (2013) report, which investigated the use of 11 intrapartum clinical indicators to compare performance of maternity services across NHS trusts in England. It applied similar sample selection to the current study which excluded multiple, pre-term and non-cephalic deliveries as well as women aged under 15 or over 45 years. It included women of all parities; however the results were stratified by parity. The sample selection in the RCOG (2013) report was defended on the basis that the care for this group of women was most affected by clinical uncertainty and varied greatly between providers. The sample selection in another study (Bragg et al. 2010), investigating variations in CS rates among English NHS trusts, was for women aged between 15-44 years who had a singleton birth, but included women of all parities and pre-term deliveries.

Maternal age, gestational age, birth weight, parity and certain clinical risk factors (10 in total based on ICD-9 codes amongst which placenta previa, hypertension, diabetes, pre-eclampsia, prior CS, etc.) were also used in a study by Peaceman et al. (2002) of risk-adjustment of caesarean delivery rates for quality improvement as they were seen as contributing to the maternal caesarean delivery risk.

Therefore, the sample selection and maternal demographic and non-clinical risk factors in the current study were similar to the studies mentioned earlier, while maternal clinical risk was uniquely derived (see Background variables) and based on NICE intrapartum clinical guidelines (NICE 2007). The socio-demographic factors, birth weight and gestational age may have been more relevant in predicting caesarean delivery but they were also used in the models for instrumental delivery and normal birth, so that the results could be compared. Each of these modes of birth may have had additional risk factors.

7.3.3 STAFFING AND TRUSTS CHARACTERISTICS AND OUTCOMES

Empirical studies investigating the effect of maternity staffing on birth outcomes are sparse. The majority of the staffing/outcomes literature relates to nurses in the acute sector. Therefore comparison of the results from this study with other studies was limited.

This study found significant associations between certain maternity staff groups and indicators. In particular, the increased chances of normal birth for all women and even greater chance for low risk women (nulliparous with singleton, live, at term birth) with increased levels of midwives. For all women, irrespective of risk, giving birth in a trust with more consultants increased the odds of instrumental delivery; while for high risk women giving birth in a trust with more doctors increased the chance of emergency CS.

7.3.3.1 CONSULTANTS O&G

For all women, irrespective of risk, giving birth in a trust with more consultants increased the odds of instrumental delivery. This finding contrasts with a study by Joyce et al. (2002, reviewed earlier) which suggested that staffing levels appeared unrelated to instrumental vaginal delivery rates and that variations in instrumental delivery rate between maternity units in their study were most significantly explained by socio-demographic factors.

This finding of a relationship between the number of consultants and the odds of instrumental delivery supports the possibility that having more consultants could lead to greater success in and therefore greater use of instrumental deliveries, which would be beneficial if it leads to reduction in the rates of caesarean section. The current study could not establish this link as instrumental delivery and emergency CS were considered independently.

Some trusts may have more consultants because of policies in place for reducing unnecessary caesarean section rates, particularly for first time mothers, to supervise

junior obstetric staff including in the use of instrumental delivery. The RCOG (2013) report maintained that all maternity units should implement the RCOG (2009b) recommendations that consultants on-call need to be present to supervise inexperienced trainees in performing operative vaginal delivery.

Senior supervision has been shown to reduce the intrapartum caesarean section rate (Althabe et al. 2004 in Loudon et al. 2010). Training and experience were seen as having an impact on clinical behaviour and the expectation was that improving these would improve the confidence to manage a complicated labour and delivery (Loudon et al. 2010). There was a call for better training in the use of instruments for vaginal operative delivery in the UK (Spencer et al. 2006 in Loudon et al. 2010). The recommendations from the RCOG in the form of Green Top guidelines (2011a) for operative vaginal delivery stated that all operators should undergo training before progressing to unsupervised use of an instrument. Loudon et al. (2010) also observed that a change in protocol in one unit, aiming to increase the senior supervision may have contributed to the reductions in failure rate in the use of conventional vacuum cups and caesarean section rate; obstetric consultants were providing increased hours of direct delivery suite cover in that unit.

There were some limitations in the dataset regarding drawing conclusions about maternity ward practice from this relationship.

This study did not examine the obstetric skill mix (consultants/doctors/trainees). There were no data for the proportion of the time spent by obstetricians on the labour ward. The outcomes of the instrumental deliveries were also unknown, for example whether they resulted in a healthy mother/baby or in undesirable events such as $3^{rd}/4^{th}$ degree tear or haemorrhage. In this study, emergency CS and instrumental delivery were treated as process indicators. Physicians' decisions to perform them will depend on the NICE clinical guidelines and the experience and confidence of the clinicians, but the actual process of decision making, the timing and the people involved could not be recorded in administrative datasets.

While skill, education, experience, organization, and leadership have an effect on the effectiveness of professional performance, they are difficult to assess (Kane et al. 2007 in the case of nurses). Usually work was done in the opposite direction, inferring skill from outcomes after other factors have been accounted for. The RCOG (2013) report found wide variations in adjusted instrumental delivery rates between hospitals in England. There were also wide variations in the instruments used (the ratio of vacuum extraction to forceps deliveries). The authors suggested that this variation may be explained by a lack of precise recommendations in the existing clinical guidelines related to the choice of instrument and inconsistencies in training within hospitals.

7.3.3.2 DOCTORS

Another significant result from the current study is that higher risk women were more likely to have increased odds of emergency CS when more doctors were available. This result partially supports the results from the study by Joyce et al. (2002) which established that the level of junior obstetric but not consultant medical staff was positively correlated with caesarean section rates. These conclusions however relate to the overall CS rates; in addition, the aggregated group of doctors in the current study could not distinguish between experienced doctors, junior doctors or trainees. NICE (2011) guidelines suggested that involvement of consultant obstetricians in the decision making for CS may reduce CS rates.

The results in the current study could have a variety of explanations. For example, doctors may judge clinical risk (particularly in relation to high risk women) and therefore make the decision to perform emergency CS without a prior attempt at instrumental delivery. Alternatively, doctors may not be confident at performing instrumental deliveries, and resort to emergency CS. It may be that not enough consultant obstetricians were available for supervision and decision making, by-passing instrumental deliveries. Or emergency CS may be a result of failed instrumental delivery.

The RCOG (2013) report found higher than expected rate of failed instrumental deliveries in 11% of the hospitals with maternity services in England, which they said may have been related to inadequate training in the use of instruments; the willingness to attempt instrumental delivery as opposed to direct emergency CS referral; or poor selection criteria for trial of forceps or vacuum in the second stage of labour. The RCOG (2013) report suggested that the rates of instrumental delivery (particularly if failed) and emergency CS should be monitored simultaneously, which this thesis analyses support.

Loudon et al. (2010) examined the trend in operative deliveries at full dilation for the period 1992-2001 and established that direct CS rates without an attempt at instrumental delivery have increased, as well as CSs in second stage of labour and at full dilatation due to both instrumental failure and reluctance to use instruments. Although the study examined a trend more than 12 years ago, their further observations revealed that the rate of full dilatation caesarean continued to increase post 2001. The authors were not certain whether these were related to reduced junior doctor hours and, therefore, clinical training and experience.

The aggregated group of doctors in this study could not distinguish between experienced doctors and trainees and therefore a measure of skill mix may have introduced some clarity or may have changed the result.

There is a range of factors which this study did not account for. These were shown to impact on the likelihood of caesarean section, such as: electronic fetal monitoring (Alfirevic et al. 2013) which may increase the likelihood of CS; while the recommendations from NICE (2011) guidelines suggested that the use of partograms, continuous support during labour from women with or without prior training, induction of labour beyond 41 weeks; labour and delivery guidelines may reduce CS rates (NICE 2011). It was also recognised that the definition of emergency CS (resulting in differences in interpretation and coding) may have contributed to the observed variations in emergency CS rates across maternity units in England (RCOG 2013) as the definition of urgency covered a wide range of clinical situations.

According to NICE (2011) guidance the classification of urgency relates to immediate threat to the life of the woman or fetus; maternal or fetal compromise which is not immediately life-threatening; no maternal or fetal compromise but needs early delivery and delivery timed to suit woman or staff.

7.3.3.3 MIDWIVES

This study found that more midwives improved all women's chances irrespective of risk of having a normal birth; this effect was even stronger for low risk women.

The study by Joyce et al. (2002) did not establish any association between midwifery staff and instrumental delivery or caesarean section after adjusting for confounders. The study did not consider normal birth or spontaneous vaginal delivery. The definition of normal birth[108] in the current study (which excluded both pre- and post-delivery anaesthetic) was stricter compared to the normal birth definition from the Maternity Care Working Party (MCWP 2007).

Midwives are expected to be the main care-giver for normal labour and birth and work in partnerships with obstetricians, anaesthetists and paediatricians in care for women with complex labours (*Safer Childbirth* 2007; *Midwifery 2020*). They are expected to be the lead professional for all healthy women with straightforward pregnancies and the key coordinator of care for women with complex pregnancies (*Safer Childbirth* 2007; *Midwifery 2020*).

Despite the positive and significant association between normal birth and levels of midwifery staffing, it is unclear exactly how these levels influenced this outcome.

It was not known how midwives were deployed within maternity services, what were the institutional policies and the culture of the work place, what were their skills and

[108] Originally suggested by BirthChoiceUK.

attitudes to labouring women and their relationship with other staff in a multidisciplinary team.

It is possible that trusts with higher levels of midwifery staffing are better able to offer continuous, one-to-one intrapartum support and midwife-led continuity models of care for which evidence suggests increase women's chances of spontaneous vaginal birth, and reduce likelihood of instrumental vaginal and/or caesarean birth (Hodnett et al. 2013; Sandall et al. 2013[109]).

This study was not able to control for models of care, neither for practice settings such as midwife-led units[110]. However, there was some indication that women giving birth in trusts with both obstetric and freestanding midwifery units may have improved chances of normal birth. Whether this was due to the exclusive midwifery care for low risk women in the freestanding midwifery units is not clear.

Continuous support during labour was recommended by NICE (2007) guideline and was seen as being *"the norm, rather than the exception"* (Hodnett et al. 2013:16). However there was evidence (Hodnett et al. 2013) of limited effectiveness of members of the hospital staff to provide continuous support; the support was more effective when it was provided by caregivers who were not employees of an institution. The authors suggested that the effectiveness of continuous intrapartum support may be improved or reduced by policies and practices in the birth setting[111]; by the nature of the relationship between the provider and labouring women; and may vary by provider characteristics (self-selection, additional duties beside labour support, limitations of institutional policies and routine practices). There was a

[109] For midwife-led continuity models of care there were no differences in caesarean birth rates.

[110] It was suggested that practice settings could be a confounding factor for the outcomes of midwife-led continuity of care (Brocklehurst 2011 in Sandall et al. 2013)

[111] Routine use of electronic fetal monitoring (EFM); availability of epidural analgesia and policies about the presence of additional support people of the woman's own choosing

warning addressed to policy makers and administrators that continuous support by midwives may not be enough in reducing caesarean rates unless other changes to policies and routines are implemented and that hospital staff may not be able to offer the same support to labouring women as non-staff members, in *"the absence of fundamental changes in the organisation and delivery of maternity care"* (Hodnett et al. 2013:16).

Midwife-led continuity of care on the other hand has been defined as care where *"the midwife is the lead professional in the planning, organisation and delivery of care given to a woman from initial booking to the postnatal period"* (RCOG 2001 in Sandall et al. 2013:3). The philosophy behind midwife-led continuity models is normality, continuity of care during labour by a known midwife (Sandall et al. 2013:2). The midwife-led continuity of care can be provided by a team of midwives or through caseload midwifery who work in a multi-disciplinary network (through consultation and referral). Based on the results (including increased chance of spontaneous vaginal birth and reduced likelihood of instrumental birth) from a Cochrane review (Sandall et al. 2013) the authors suggested that *"most women should be offered midwife-led continuity models of care"* (p.2) but caution should be applied in offering that advice to women with medical or obstetric complications. Other factors may have affected the likelihood of spontaneous vaginal birth for women randomised to midwife-led continuity of care; these included attendance at birth from a known midwife, increased mobility, and philosophy of care (Sandall et al. 2013). Despite recognising that philosophy of care was difficult to isolate, there was an indication that it may have had a separate effect to the birth setting (in 9 trials care was provided on the labour ward, i.e. the benefits from that care were evident even when midwives provided intrapartum care in hospital setting). The review recommended midwife-led continuity of care to the policy makers aiming at normalising birth and suggested considering the financial implications for achieving midwife-led care.

Overall, future work investigating the relationship between midwifery staffing levels and birth outcomes should incorporate at least models of care in the causal pathway.

This study did not find any statistically significant relationship between healthcare assistants and any of the indicators. This does not imply that this support group can not contribute to better outcomes or experiences of birth for women.

7.3.3.4 TRUSTS CHARACTERISTICS

There were some indications that women giving birth in London trusts or in teaching trusts have reduced chances of normal birth and that women giving birth in trusts with both obstetric and freestanding midwifery units may have improved chances of normal birth. Foundation status of the trusts seemed to have no impact on the outcomes.

7.3.3.5 CONCLUSION

This study found an important relationship between the level of instrumental delivery and numbers of consultants, and between the likelihood of emergency CS and numbers of doctors, and between numbers of midwives and the odds of normal birth. These findings have clinical and policy implications, discussed below.

Overall, however, there were very few significant relationships between the three indicators and staffing, and almost no relationship between the selected trusts characteristics and the indicators.

These few relationships may have been due to the small initial variations in staffing levels between trusts, resulting possibly from following guidelines on staffing requirements. In addition, aggregated staffing data may have contributed to the inability to explain variations, and reduced sensitivity to establish the effect of staffing on the outcomes. If workforce and patient level data were available at obstetric/midwifery unit level of analyses, this may have changed the results (in terms of variations and the association between staffing and the outcomes).

The rather few relationships does not imply that having additional staff may not achieve better intrapartum outcomes for women. Additional measures of the process

and outcomes of maternity care are needed before one can confidently comment on the effects of the levels of obstetric, midwifery and supporting staff in relation to the investigated outcomes.

7.3.4 WOMEN'S CHARACTERISTICS AND CLINICAL RISK

The results confirmed a strong and significant relationship between women's characteristics, clinical risk and the three indicators. Overall age, clinical risk, birth weight and gestational age all had significant and strong relationships with the outcomes. Ethnicity and area of deprivation (IMD) though significant had weaker impact. Clinical risk was the strongest predictor for emergency CS and normal birth, while gestational age was the strongest predictor of instrumental delivery.

Maternal age was a strong predictor for all the three outcomes. The odds of emergency CS increased in each older age group compared to 15-19 years old. The results in this study also showed a positive curvilinear (inverted-U) relationship between maternal age and instrumental delivery (reaching a peak in age group 30-34 and declining afterwards); and a negative relationship between age and normal birth. The study by Main et al. (2006) found that from the age of 17 (cut-off at 41) there was nearly a straight-line increase in caesarean births for their sample of NTSV (nulliparous, at term, singleton, vertex). Other studies (Paranjothy et al. 2005; Peaceman et al. 2002; Bragg et al. 2010) have also found that the odds of CS increased with age. However the samples in these studies are not comparable to this study which only looks at nulliparous women; the studies also investigated the odds of overall CS or the odds of CS before or during labour (Paranjothy et al. 2005) and not explicitly emergency CS. Essex et al. (2013) established that for primiparous women, operative births rose steeply with increasing maternal age, including instrumental delivery. The results in this study showed a positive curvilinear (inverted-U) relationship between maternal age and instrumental delivery (reaching a peak in age group 30-34 compared to 15-19, and declining afterwards)

In this study, ethnicity was a much stronger predictor of instrumental delivery and less so for emergency CS and normal birth. Overall Asian women had slightly higher odds of instrumental delivery and emergency CS or lower odds of normal birth compared to White women. Afro-Caribbean women had lower odds of instrumental delivery, but higher odds of emergency CS or normal birth compared to White women. The study by Essex et al. (2013) has also established that for first-time mothers, Black women were less likely to have an instrumental vaginal birth but Pakistani or Bangladeshi women were less likely to have emergency caesarean section compared with White women.

Essex et al. (2013) established that ethnic minority[112] women were at an increased risk of emergency caesarean section using data from the Millennium Cohort Study (MCS)[113]. Bragg et al. (2010) have also found increased odds of CS for Afro-Caribbean women, though this related to the overall CS and their sample included women of all parity and pre-term deliveries. Paranjothy et al. (2005) found that women from ethnic minority groups had lower odds of CS before labour, and increased odds of CS in labour with a possible explanation from other studies (Petrou et al. 2001 in Paranjothy et al. 2005) that women from these groups may not be accessing antenatal care; have less antenatal visits; or late antenatal booking (after 18 weeks of gestation). The consequences of that were that many of the problems in their pregnancies were diagnosed later and some of them possibly even after the onset of labour.

Areas of deprivation measured with IMD in this study had a significant but weak relationship with the three outcomes. Women in most deprived areas had increased odds of emergency CS or normal birth and reduced odds of instrumental delivery compared to women from least deprived areas after adjusting for the effects of other

[112] 'Minority ethnic' wards were those in which at least 30% of the total population were in the categories 'Black' or 'Asian'.

[113] A large-scale study of babies born between September 2000 and January 2002

predictors. Bragg et al. (2010) did not find any differences in CS rates by area of deprivation. Pregnant women of lower socio-economic status were at higher risk of poorer infant outcomes in relation to low birth weight and pre-term birth (Macfarlane and Mugford 2000). They may have poorer diet, are more likely to be obese and be involved in risky behaviour (Dowd 2007). UK studies have shown that women of lower socio-economic status may have insufficient knowledge and restricted choices about their care (Bowes and Domokos 2003), and they may be more willing to accept interventions in labour (Green and Baston 2007). Socio-economic status is not the same as area of deprivation. Area based measures could not accurately reflect women's individual socio-economic status. Though area of deprivation (IMD) may be useful to control for inequalities in health, it also cannot directly determine whether the differences across areas are due to characteristics of the areas or due to differences in health and types of individuals living in these areas (Diez Roux 2001).

Both emergency CS and normal birth were most predicted by NICE clinical risk, with high risk women more likely to have emergency CS, while low risk women were more likely to have a normal birth. NICE clinical risk was a much weaker predictor for instrumental delivery. Given that NICE (2007) clinical risk was a composite measure it was difficult to compare the relative impact of the clinical conditions with the impact of clinical factors in other studies. In the study by Bragg et al. (2010) clinical factors including diabetes, hypertension and placental problems predisposed women to having a caesarean. For the clinical risk composite measure in this study, it was acknowledged that each included condition may have a 'spectrum of risk', and be associated with different adverse outcomes and that assignment of risk was affected by incomplete or inaccurate coding thus trusts with poor coding may present with a higher proportion of lower risk women. Despite these limitations the authors of the composite risk measure used in this study were confident that the risk assignment tool distinguished well between the groups of women.

Gestational age was the strongest predictor for instrumental delivery and second strongest for normal birth. Women with gestational age greater than 42 weeks had higher odds of emergency CS or instrumental delivery and reduced odds of normal

birth compared to women with gestational age between 37 and 41 weeks. Birth weight was the second strongest predictor (after NICE clinical risk) for emergency CS but a weaker predictor for instrumental delivery. Both groups of women with small (<2500g) and big (>4500g) babies had higher odds of emergency CS compared to women with babies weighing between 2500g and 4500g; this relationship was reversed for instrumental delivery. Birth weight was a relatively strong predictor for normal birth, where women with big babies had reduced odds, while women with small babies had higher odds of normal birth.

It was clear that case-mix adjustment was necessary to allow for case-mix differences across trusts and to try and isolate the effect of staffing on the outcomes. Ultimately adjustment is always limited because of the availability, relevance or error in measurement of the predictive factors, particularly when using routinely collected data; but also by the incomplete knowledge of all the factors that may influence an outcome (Weir 2004). Even if the control was close to ideal, there is a limit of how much of the variations in outcomes could be explained by statistical models (Weir 2004). Conclusions therefore, when a single and widely accepted adjustment model does not exist, could vary (Iezzoni 2003).

The two sensitivity analyses which were performed by 1) including the missing data as an extra level in each of the categorical explanatory Level-1 predictors; and 2) exploring the cross-level interaction effects between staffing ratios and clinical risk; revealed in both analyses that the strength or the direction of the relationships between women's characteristics and the three outcomes did not change much.

Some of the missing categories coefficients though were significant and the inclusion of the missing data revealed a significant and negative relationship between midwives and emergency CS; and a positive significant relationship between trusts with both OU and FMU configuration and normal birth. Missing data analyses therefore showed that most of the results remained stable with few differences, particularly the effect of having more midwives in reducing the odds of emergency CS.

7.3.5 VARIATIONS IN OUTCOMES

Despite the significant independent effect of staffing on some of the indicators, this study found that only a small percentage (1%-2%) of the total variability in emergency CS, instrumental delivery and normal birth was attributable to variations between trusts.

The very low initial (null models) shares of between trusts variations in emergency CS, instrumental delivery and normal birth were not explained much by women's characteristics, staffing levels or trusts characteristics. In fact adding NICE clinical risk and staffing variables increased slightly the trust variance in emergency CS model; adding gestational age, birth weight and trusts characteristics increased slightly the trust variance in instrumental delivery model; and adding NICE clinical risk increased the trust variance in normal birth model. It was suggested that adding individual level variables with strong effects tend to increase trust-level variances or that variance increase may be observed if the newly added variables explain mainly within trusts variations (Snijders and Bosker 1999).

The selected sample of nulliparous women was relatively homogeneous, and even though this may have been better suited to understand the non-clinical factors relationship with birth outcomes, it may also have dampened other significant variations due to parity, premature births or previous obstetric history for example. Stivanello et al. (2013) showed that models with more risk adjustment variables explained more variation (in CS rates) and when most important factors (in their case for caesarean delivery) were omitted and thus a more homogenious population was investigated, the predictive power of the models was poorer but the observed to expected ratios (of CS rates) were similar across models. The predictive powers of the models in the current study were also poor.

The results may have been affected by the sample size, mainly at the highest unit of analysis (Institute for Digital Research and Education, UCLA online[114]); in this study the number of NHS trusts with maternity services was 143 for emergency CS and instrumental delivery models and 127 for Normal birth model. This may have impacted on the stability of the trusts effect estimates as well as the between trusts variations. Possibly having less individual level observations in Level-1 and more trusts (maternity units) in Level-2 would have been more favourable to the research questions asked in this study.

The results may have also been affected by the use of quasi-likelihood approach to estimating the model parameters. According to the methodological literature *"there are no closed form solutions"* for Generalised Linear Mixed Models, and a quasi-likelihood approach[115] (via a Taylor series expansion) is one way to approximate the likelihood (IDRE, UCLA online). GLMM procedure in IBM SPSS 22 offers only the quasi-likelihood procedure in which parameters are estimated so to maximize the quasi-likelihood. These estimates though are not the true maximum likelihood estimates. The advantages of the quasi-likelihood approach are that it is the fastest (compared to other approaches) thus useful for initial exploration and for large datasets. There is however a bias associated with this approach and quasi-likelihoods are not recommended for final models or statistical inference (IDRE, UCLA online). The Null models were checked with STATA SE v 12 (StataCorp LP) which uses numerical integration and the results did not differ from the ones obtained using quasi-likelihood approach in IBM SPSS 22 and MLwiN 2.26. The results may need to be confirmed with the Markov chain Monte Carlo (MCMC) algorithm available in MLwiN 2.26 which was not explored.

---

[114] http://www.ats.ucla.edu/stat/stata/dae/melogit.htm

[115] Other approaches include numerical integration (to approximate the true likelihood) or Monte Carlo methods (including Metropolis-Hastings algorithm and Gibbs sampling which are types of Markov chain Monte Carlo (MCMC) algorithms), http://www.ats.ucla.edu/stat/stata/dae/melogit.htm.

7.4 CONTRIBUTION TO KNOWLEDGE

This study emerged from the realisation that a large body of international research evidence suggested a strong link between nurse staffing and patient outcomes and that very little research was done outside acute general service sector, including maternity. At the same time issues that drove the expansion of nurse staffing research, such as shortages of nurses; demographic changes in the population; patent-centeredness; cost reductions and patients' safety were relevant to maternity services as well. Given that the literature on the relationships between staffing and outcomes in maternity was sparse, this study contributed to better understanding of these relationships and the limitations involved. It also confirmed the importance of women demographic and clinical risk characteristics on their birth outcomes.

The study used a comprehensive composite measure of clinical risk which was based on the NICE 2007 intrapartum care guidelines. The results confirmed the importance of considering risk adjustment (particularly age and clinical risk factors), when comparing trusts outcomes.

The results highlighted some significant associations between maternity staffing and selected intrapartum outcomes, the conclusions and recommendations though are limited by the retrospective, cross-sectional study design and are generalisable to women with similar demographic and obstetric characteristics as the studied population (women aged 15-44, who were nulliparous and had a term, singleton, live birth).

A main contribution of this study was in applying multilevel logistic analysis by simultaneously including both trusts- and individual-level predictors in regression equations with women as the unit of analysis. This approach allowed for isolating maternity staffing and trusts effects after individual-level confounders have been accounted for; and also allowed for examination of women characteristics as modifiers of the staffing and trusts effects (and vice-versa) (Diez Roux 2001). Multilevel analysis also examined simultaneously the within- and between-trusts

variations in emergency CS, instrumental delivery and normal birth and the extent to which between-trusts variations in outcomes are explained by individual-level (women) and trust-level (maternity staffing, configuration) factors (Snijders and Bosker 1994).

Another contribution of this study was the inclusion of all the major maternity staffing groups (obstetricians, doctors, midwives and healthcare assistants), which was a major critique in the nurse staffing/outcomes literature, where in majority of the studies nurses were considered in isolation of other staff groups, particularly doctors. This study though could not establish whether trusts with higher midwifery staffing levels have also had more and better qualified doctors and therefore better outcomes.

7.5 STRENGTHS AND WEAKNESSES OF THE STUDY

7.5.1 GENERALISABILITY AND REPRESENTATIVENESS OF THE STUDY SAMPLE

A contribution of this study was that the analyses was applied after linking two large datasets: 1) individual-level retrospective routinely collected administrative dataset (HES) which provided information on all women who gave birth in English NHS trusts between 1st April 2010 and 31st March 2011; and 2) NHS maternity workforce census data which provided information on all medical and non-medical staff employed by the maternity services of the NHS trusts in England on 30 September 2010.

Even though the study only selected a sample of women aged 15-44, who were nulliparous and had a term (>=37 weeks), singleton, live birth, the sample was large (261,895 women) and represented nearly 40% of all women who gave birth in England during April 2010 and March 2011. Overall the sample differed from the total HES population in having a higher proportion of women in the youngest age group (15-19) and a smaller proportion of women aged greater than 35; a marginally smaller proportion of Asian and Afro-Caribbean women; a smaller proportion of women living in the most deprived area based on IMD; higher proportion of women

classified as low risk than high risk based on NICE (2007) criteria; a lower proportion of women giving birth to babies weighing below 2500g and a marginally higher proportion of women with GA greater than 41 weeks. In comparison, a higher proportion of women in the HES population were classified as NICE (2007) high risk than low risk. A higher proportion of women in the sample had interventions - emergency CS (19.3% vs 14.7%) and instrumental delivery (21.8% vs 12.4%); and a lower proportion of women had normal birth (24.7% vs 33.6%).

The current study results will therefore be only generalisable to a population of women with similar demographic and obstetric characteristics.

7.5.2 VARIATIONS IN OUTCOMES

The partitioning of the residual variance in the outcomes due to the individuals and trusts is the essence of multilevel modelling approach, and one of the reasons for applying such analyses was to quantify the significance of trusts characteristics and staffing in understanding individual outcomes (in this case emergency CS, instrumental delivery and normal birth). The variance component is usually quantified through the calculation of the variance partition coefficient (VPC) and is interpreted as the proportion of the total variation attributable to variation between trusts. This though is not the same as variation between trusts.

When the response variable is continuous there is a clear distinction between individual level variance and trust level variance. When the response variables are binary as in the current study this distinction is not clear (Merlo et al 2006), which makes the calculation and interpretations of VPC difficult (see 4.10.7 in Methods chapter for alternative approach in calculating VPC for binary response). This is because of the non-linear (logit) relationship between the covariates and the response variable. In addition, in multilevel logistic regression, the individual level variance and trust level variance are not directly comparable, because the individual level variance is on the probability scale (and depends on the probability of the outcome), while the trust level variance is on the logistic scale. Larsen and Merlo (2005) have

introduced a measure which translates trust level variance into odds ratio scale, which makes it directly comparable with the odds ratios for individual or trust level variables. They called that measure the Median Odds Ratio (MOR). MOR and VPC are closely related but they differ in that VPC is a function of both trust and individual level variances, while MOR is a function only of the trust level variance. In short they have provided a formula[116] to calculate MOR (Larsen and Merlo 2005), which for example establishes the following relationships: If the trust level variance is small (for example it was i.e. σ^2=0.044 for emergency CS in this study), this corresponds to VPC=1.32% and MOR=1.22. Though MOR and VPC are very close in this example with small variance they measure different aspects.

MOR quantifies the variation between trusts by comparing two women with the same characteristics from two randomly chosen different trusts. The MOR ratios in the final three models (Tables 5-11; 5-19 and 5-25 in Chapter 5) were low odds ratios[117], suggesting that there were little variations between trusts in the probability of emergency CS, instrumental delivery or normal birth for the selected sample of nulliparous women. If MOR=1 there would be no difference in the probability of emergency CS/instrumental delivery/normal birth between trusts. In the median case for example the residual heterogeneity between trusts increased by 1.22 times the individual odds of having an emergency CS when randomly picking out two women in different trusts; that is if a woman moved to another trust with a higher probability of emergency CS, her risk of having an emergency CS would (in median) increase 1.22 times.

Also the residual heterogeneity between trusts (MOR=1.22) for example was of lesser relevance than the impact of maternal age (OR=2.169 for women aged 30-34)

---

[116] (MOR≈exp(0.95$\sqrt{\sigma^2}$)

[117] Emergency CS (trust variance=0.044; MOR=1.22; VPC=1.32%); Instrumental delivery (trust variance=0.053; MOR=1.24; VPC=1.59%); Normal birth (trust variance=0.055; MOR=1.25; VPC=1.64%)

or clinical risk (OR=2.325 for high risk women) or gestational age (OR=1.638 for women with gestational age >41 weeks) in understanding the odds of emergency CS.

7.5.3 LIMITATIONS OF THE DATA

The data did not allow controlling for complications during labour. Clinical risk (NICE 2007 intrapartum care definition) in this study was determined retrospectively from diagnoses at the antenatal stage and at onset of labour. The clinical conditions NICE (2007) list were used to facilitate the categorization of high risk women, which in turn was meant to determine suitability for care in obstetric unit. A retrospective study by Sinclair et al. (2001) on risk assessment of women (2200 women in Northern Ireland) suitable for midwifery care found that a third of the women were low risk at the onset of labour, but 20% of them moved to high-risk category during first, second and third stages of labour. The main reasons for change in risk status were augmentation of labour and meconium liquor. Nulliparous women were at most risk of developing complications. Still 90% of all women in the sample had a vaginal birth regardless of risk. The *Birthplace* in England (2011) prospective cohort study which focused on low risk women (NICE 2007 intrapartum care definition) also established that 20% of those low risk women, whose planned place of birth at the start of labour was obstetric unit, presented with at least one complication at the start of care in labour (compared to fewer than 7% for all other non-obstetric planned places of birth). The most common complications identified by the attending midwife were meconium stained liquor and prolonged rupture of membranes (>18 hours); other included: protenuria, hypertension, abnormal vaginal bleeding, non-cephalic presentation and abnormal fetal heart rate.

The models in this study could not account for obesity, smoking, drinking, diet, stress, mothers' preferences, health care professionals beliefs and attitudes, their training and experiences, many organisational factors and models of care. Stivanello et al. (2013) found that as the number of risk adjustment variables in their caesarean delivery models increased, so did the amount of variation explained, but that simpler

adjustment worked as well as complex ones. However obesity and smoking would have been important to include in the models of the current study.

Data for smoking exists in HES but was of poor quality (not well recorded at trust level). In the general population, a trend of reduced alcohol consumption[118] and smoking[119] between 1998 and 2009 was also observed.

Obesity is known to contribute to higher risks of heart disease, high blood pressure and diabetes and to additional risks for pregnant and labouring women such as macrosomia, which translates into increased demand for extra resources (including staff) during the antenatal and labour stages. They may need increased levels of maternal and fetal monitoring, more surveillance and screening (CMACE and RCOG 2010). Obese women are at higher risk of intrapartum complications, including anaesthesia and caesarean section. This will demand senior obstetric and anaesthetic involvement and an antenatal anaesthetic assessment, and potential for emergency caesarean section (joint CMACE and RCOG 2010). Body mass index can also influence failed instrumentation (Gopalani et al. 2004 in Loudon et al. 2010), and caesarean section rate (Kabiru and Raynor 2004 in Loudon et al. 2010). NICE (2007) recommends that women with BMI greater than 35 should give birth in a consultant-led obstetric unit with neonatal services. The UK prevalence of women with a known BMI ≥35 at any point in pregnancy, who give birth at 24+ weeks' gestation, was 4.99% (approximately 38,500 maternities each year in UK) while there were 2.01% of pregnant women with BMI greater than 40 (CMACE 2010).

[118]http://www.statistics.gov.uk/downloads/theme_health/alcohol-related-deaths-in-the-uk-1991-2009.xls

[119]http://www.statistics.gov.uk/downloads/theme_compendia/GLF09/GeneralLifestyleSurvey2009.pdf

Further data on gestational age (16.8%) and birth weight (10.2%) in HES, though getting better in quality in recent years had many missing values. There were 7 trusts with 100% missing data on birth weight and 10 trusts with 100% missing data on gestational age. There was much less but still missing data on maternal age (0.3%), ethnicity (4.5%), area of deprivation (0.8%), multiplicity (0.1%), delivery method (0.1%) and live births (11.2%). The sensitivity analyses, which included the missing data as an additional category in the predictor variables, showed overall stability of the estimates and direction of the relationship between the predictors and the three indicators, although certain small differences were identified. Imputation of missing data was not attempted.

The workforce data was an annual census data therefore it was not possible to control for daily, weekly or monthly fluctuations in staffing levels. Availability of trust data on suspension of maternity services (yes/no or number of times within a year) due to staff shortages or insufficient beds could possibly be used as a proxy for staffing fluctuations or demand pressures on maternity services within trusts, which may have had additional impact on the variations in outcomes. That data exists for London trusts in 2011/12 (from London LSA Annual report 2011/12).

The indicators should really be more directly linked to the staffing group or unit which had most responsibility over aspects of care measured by the outcomes, while the provision of care by multidisciplinary teams could make it difficult in practice to link indicators to exact staffing groups (Johal et al. 2013).

Staffing was a trust level variable because of the data availability, therefore the quantity of staff FTE provided to each woman in labour or indeed which staff groups assisted at birth or what was the skill mix were unknown.

The FTE/birth staffing ratios were based on the total staff FTE numbers in each NHS trust with maternity services. It was not clear how these staff groups were deployed and what proportion of their time was spent on the labour ward (of interest because all the investigated outcomes were intrapartum). It was not clear how much time

consultants O&G spent in gynaecology and how much in obstetrics and on the labour ward. The same applied to doctors' time spent on the labour wards, while midwives apart from deployment in labour wards are also deployed in the antenatal and postnatal wards as well as in the community. It was also not possible to distinguish between midwives deployed within the obstetric/alongside or freestanding midwifery units, neither whether women delivered in an obstetric, alongside midwifery or freestanding midwifery units. In addition to not knowing how the multidisciplinary teams were organised, there was no information on how they interacted or communicated with each other, as joint effects of staffing and elements of practice environment on outcomes exist (Aiken et al. 2002b). Another study from the UK (Griffiths et al. 2011) investigating the relationship between nurse staffing and outcomes in primary care found that the strongest predictors of quality of clinical care were not staffing levels but the organisational factors (education and training and use of patient experience surveys).

Staffing levels overall are considered a blunt instrument to measure quality of care (Weir 2004). They should really be investigated together with the main effort (the process of care) and the outcomes of that care.

Finally, Sinha et al. (2012:3) warns that "*traditionally, 'quality' is equated with 'non-susceptibility to bias' and this poses particular problems for retrospective studies using administrative data"* (Sinha et al. 2012:3). Strength of the retrospective methodology is seen in the potential to apply more accurate risk assessment criteria (Sinclair et al. 2001), while on the other side because of the retrospective nature of HES there was a risk of a post-hoc justification of using certain procedures (Johal et al. 2013).

7.6 POLICY AND CLINICAL IMPLICATIONS

This study found that having more midwives will improve the chances of normal birth for all women (aged 15-44, who were nulliparous and had a term (>=37 weeks), singleton, live birth). The impact is even higher for similar low risk women.

However further research on the actual mechanisms by which midwifery care affects normal birth outcome is needed. These results could also not inform workload planners on how and where to deploy more midwives. Given that the study was limited to intrapartum care outcomes for women, no conclusions could be drawn on what the impact of an increase in midwifery workforce could be on outcomes for women and their babies on the antenatal and postnatal wards, neither in the community settings.

Even if we assume that a casual relationship exists between midwifery staffing and normal birth, policy and financial impediments currently exist in increasing the midwifery workforce. *"Stakeholders in England believe more could be delivered for less money with better outcomes if there were more midwife-led birth centres available"* (House of Commons Committee 2014:5). For this to be achieved more midwives need to be employed. Currently there is a national shortage of 2,300 midwives necessary to meet current birth rates (House of Commons Committee 2014). The available funding for many maternity services may not be sufficient to employ enough midwives and consultants to provide high quality, safe care. There are indications that the new tariff payments for maternity care introduced in April 2013 may not provide enough income to providers to ensure high quality, safe care and was seen as holding back the provision of more birth centres (House of Commons Committee 2014). The new tariff payments aimed to compensate for the additional cost of care for high risk women but interventions as such are not rewarded. In addition defining the best level of midwifery staffing in order to achieve high quality, safe care will require cost-effectiveness analysis (Kane et al. 2007). This will possibly mean comparing the current midwives/birth ratios with estimated changes in midwifery staffing needed to achieve a desirable patient outcome (i.e. normal birth). That will require evidence for a causal relationship between midwifery staffing levels and normal birth. The current cross-sectional study cannot provide that evidence. Therefore in the absence of randomised controlled trials, future work will require longitudinal design.

This study established positive association between consultants staffing levels and instrumental delivery. Whether the outcomes of these instrumental deliveries are healthy mother and baby and reduced primary caesarean rates need further investigation. There are tangible healthcare costs, but also health and future pregnancies implications for first-time mothers with reduction in primary caesarean rates.

This research showed that for nulliparous women, emergency CS, instrumental delivery and normal birth were largely determined by their characteristics and clinical risk. This has important policy implications when planning maternity care for first-time mothers, as well as for promoting lifestyle changes and providing women with additional information on the impacts of age and clinical risk on their outcomes. There is evidence in this research to support the life-course approach to women's healthcare advocated by the RCOG (2011) *High Quality Women's Health Care* by promoting health and lifestyle in every interaction a woman has with the health service, at every age rather than the constant fight against disease and ill health which are burden to the financial and human resources in the current health system.

HES data quality is improving but some trusts are still presenting with 100 per cent missing data on gestational age and birth weight. In addition data on body mass index and smoking is either missing or of poor quality. Availability of disaggregated staffing data at maternity unit level may also be beneficial in future research. Therefore any future data quality improvement initiatives should be supported.

7.7 FUTURE WORK

This was an exploratory study, given that very little was known about the relationship between maternity staffing and birth outcomes. As such it probably raised more questions than it answered. It became clear that more disaggregated workforce data and more complete and accurate individual-level data may reveal more about these relationships. Future work will need better case-mix adjustment (to include smoking and obesity), as well as including models of care in the causal

pathway. Staffing data was measured at trust level; maternity unit level measurement may have revealed more variations; but ultimately even unit-level staff measurements will not capture clinician/patient interactions nor the variations in midwives/obstetricians resources distribution due to individual women complex needs.

This study did not aim to evaluate the selected indicators as measures of quality of care but if they were to be used as measures of quality, they should really be linked with neonatal outcome measures and other maternal outcomes (such as satisfaction with care), which result from the process of care (interventions or models of care) (Main et al. 2006). Ultimately the relationship between indicators which measure structure, process and outcome should be investigated to gain better understanding of how to improve quality of maternity care. This can be achieved through longitudinal study design, as randomised trials may not be feasible.

Finally if normal birth is to be treated as a measure of quality of midwifery-led care and a positive outcome of that care which can be achieved through higher midwifery staffing levels, a measure of the process of that care should be included in the causal pathway as well, such as one-to-one care or continuous support during labour, which have been identified as beneficial in achieving normal birth.

REFERENCE

AbouZahr C. (2003) *Global burden of maternal death.* In: British Medical Bulletin. Pregnancy: Reducing maternal death and disability. British Council. Oxford University Press.2003. 1-13. www.bmb.oupjournals.org

Abrahms C., Davy K. (2002) *Linking HES maternity records with ONS birth records.* Health Stat Q 2002;(Spring Edition):22–30.

AHRQ (2003) *Quality Indicators, Guide to Patient Safety Indicators.* AHRQ, March 2003, Version 3.1 (March 12, 2007). Available at: http://www.premierinc.com/safety/topics/guidelines/downloads/Pt-Safety-psi-guide-v31-2007sm.pdf

AHRQ (2006) *Hospital Nurse Staffing and Quality of Care*. United States Agency for Healthcare Research and Quality, University of California San Francisco-Stanford Evidence-Based Practice Center. Available at: http://www.ahrq.gov/ research/nursestaffing/nursestaff.htm

AHRQ (2007) *Nurse Staffing and Quality of Patient Care, Structured Abstract.* March 2007. Agency for Healthcare Research and Quality, Rockville, MD. http://www.ahrq.gov/clinic/tp/nursesttp.htm

Aiken LH. et al. (2002) *Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction*. JAMA, 288 (16).1987-93

Aiken LH., Clarke SP., Sloane DM. (2002b) *Hospital staffing, organization, and quality of care: cross-national findings*. International Journal for Quality in Health Care 2002,14:5-13

Aiken LH., Clarke SP., Cheung RB., Sloane DM., Silber JH. (2003) *Educational levels of hospital nurses and surgical patient mortality*. JAMA. 2003;290:1617–1623.

Aiken L., Sloane D., Bruyneel L., Van den Heede K., Griffiths P., Busse R., et al. for the RN4CAST consortium (2014) *Nurse staffing and education and hospital mortality in nine European countries: a retrospective observational study*. The Lancet, Early Online Publication, 26 February 2014 doi:10.1016/S0140-6736(13)62631-8

Albers LL. and Borders N. (2007) *Minimizing genital tract trauma and related pain following spontaneous vaginal birth.* Journal of Midwifery and Women's Health 52(3):246-53.

Albury D. (2009) *A perfect storm: creating the conditions for radical public service innovation*. Putting Citizens and Businesses in Control presentation, 14 May www.cabinetoffice.gov.uk/media/212012/david_albury.pdf

Alfirevic Z., Devane D., and Gyte GM. (2006) *Continuous cardiotocography (CTG) as a form of electronic fetal monitoring (EFM) for fetal assessment during labour.* Cochrane Database of Systematic Reviews 2006, Issue 3. Art. No.: CD006066.

DOI: 10.1002/14651858.CD006066. Available from: www.nelh.nhs.uk/cochrane.asp

Alfirevic Z., Devane D., and Gyte GM. (2013) *Continuous cardiotocography (CTG) as a form of electronic fetal monitoring (EFM) for fetal assessment during labour.* Cochrane Database Syst Rev. 2013;5:CD006066.

Althabe F., Belizan JM., Villar J., Alexander S., Bergel E., Ramos S., et al. (2004) *Mandatory second opinion to reduce rates of unnecessary caesarean sections in Latin America: a cluster randomised controlled trial*. Lancet 363:1934–1940.

Althabe F., Sosa C., Belizan JM., Gibbons L., Jacquerioz F., Bergel L. (2006) *Caesarean section rates and maternal and neonatal mortality in low-, medium-, and high-income countries: an ecological study.* Birth 2006 Dec:33(4):270-7.

Althabe F. and Belizán JM. (2006b) *Caesarean section: the paradox.* Lancet. 2006; 368 (9546):1472-3.

Anderson DA. and Aitkin M. (1985) *Variance Component Models with Binary Response: Interviewer variability*. J.R. Statist. Soc. B 47:203-10.

Anim-Somuah M., Smyth R., Howell C. (2005) *Epidural versus non-epidural or no analgesia in labour* . Cochrane Database of Systematic Reviews (4) : CD000331.

Ashcroft B., Elstein M., et al. (2003) *Prospective semistructured observational study to identify risk attributable to staff deployment, training, and updating opportunities for midwives.* BMJ 327(7415): 584.

Ayanian, J.Z. (1999). *Using Administrative Data to Assess Healthcare Outcomes*. Eur Heart J, 1999; 20:1689-91

Aylin P., Lees T., Baker S., et al. (2007) *Descriptive study comparing routine hospital administrative data with the Vascular Society of Great Britain and Ireland's National Vascular Database*. Eur J Vasc Endovasc Surg 2007;33:461–5

Aylin P., Bottle A., Majeed A. (2007) *Use of administrative data or clinical databases as predictors of risk of death in hospital: comparison of models*. BMJ 2007;334:1044–7.

Baghurst PA. (2012) *The case for retaining severe perineal tears as an indicator of the quality of obstetric care.* Australian and New Zealand Journal of Obstetrics and Gynaecology 2013; 53: 3–8.

Bailit J., Love, Dawson N. (2006) *Quality of obstetric care and risk-adjusted primary cesarean delivery rates*. American Journal of Obstetrics and Gynecology. 2006;194:402–07.

Ball JA., Bennett B., Washbrook M., Webster F. (2003) *Birthrate Plus Programme: Further issues in deciding staffing needs.* Br J Midwifery 2003;11(7):416–19

Ball JA. and Washbrook M. (2010a) *Workforce planning in midwifery: an overview of 8 years*. British Journal of Midwifery, vol.18, N 8.

Ball JA. and Washbrook M. (2010b) Developing a real-time assessment of staffing needs in delivery suits. British Journal of Midwifery, vol.18, N 12.

Barros FC., Victora CG., Barros AJ., Santos IS., Albernaz E., Matijasevich A., et al. (2005) *The challenge of reducing neonatal mortality in middle income countries: findings from three Brazilian birth cohorts in 1982, 1993, and 2004.* Lancet 2005; 365 (9462): 847–54.

Beake S., McCourt C., and Page L. (2001) *Evaluation of one-to-one midwifery: second cohort study*. London: Centre for Midwifery Practice, Thames Valley University; 2001. Available from: http://www.wolfson.tvu.ac.uk/cmp/reports/OnetoOneReport2001.pdf

Beckman MM., Garrett AJ. (2005) *Antenatal perineal massage for reducing perineal trauma. Cochrane Database of Systematic Reviews*. 2005, Issue 1, Article No:CD005123

Belizán JM., Althabe F., Cafferata ML. (2007) *Health Consequences of the Increasing Caesarean Section Rates*. Epidemiology. 2007; 18 (4): 485-6.

Bell JS., Campbell DM., Graham WJ., Penney GC., Ryan M., Hall MH. (2001) *Can obstetric complications explain the high levels of obstetric interventions and maternity service among older women? A retrospective analysis of routinely collected data*. British Journal of Obstetrics & Gynaecology. 108(9) 2001, pp.910-918

Benjamin Y., Walsh D., Toub N. (2001) *A comparison of partnership caseload midwifery care with conventional team midwifery care: labour and birth outcomes*. Midwifery 2001; 17(3):234-40.

Betrán AP., Merialdi M., Lauer JA., Bing-Shun W., Thomas J., Van Look P., Wagner M. (2007) *Rates of caesarean section: analysis of global, regional and national estimates.* Paediatr Perinat Epidemiol. 2007; 21 (2):98-113.

Biro MA., Walderstrom U., Pannifex JH. (2000) *Team midwifery care in a tertiary level obstetric service: a randomized controlled trial.* Birth.2000 Sep; 27(3):168-73

BirthChoiceUK (2005) Normal Birth: Welcome to the BirthChoiceUK normal birth pages.

BirthChoiceUK (2013) Historical statistics. Available from: http://www.birthchoiceuk.com/Professionals/BirthChoiceUKFrame.htm?http://www.birthchoiceuk.com/Professionals/statistics.htm

Birthplace in England Collaborative Group (2011) *The Birthplace in England national prospective cohort study: perinatal and maternal outcomes by planned place of birth in 'low risk' women.* BMJ 2011; 343:d7400.

Birthplace in England Research Programme (2010) 2010 Update: Configuration of Maternity Care Unit Questionnaire.

Blake S., McCourt C., and Page L. (2001) *Evaluation of one-to-one midwifery: second cohort study.* London: Centre for Midwifery Practice, Thames Valley University; 2001. http://www.wolfson.tvu.ac.uk/cmp/reports/OnetoOneReport2001.pdf

Blegen MA. and Vaughn T. (1998) *A multisite study of nurse staffing and patient occurrences.* Nurs Econ. 1998;16:196–203.

Bliese P. (2012) *Multilevel Modeling in R (2.4). A brief Introduction to R, the multilevel package and the nlme package*. Copyright © 2012, Paul Bliese.

Bojo AKFS., Hall-Lord ML., Axelsson O., Uden G., Larsson BW. (2004) *Midwifery care: development of an instrument to measure quality based on the WHO's classification of care in normal birth.* Journal of clinical nursing, 13, 75-83.

Bolker B., Brooks M., Clark C., Geange S., Poulsen J., Stevens H., White JS. (2009) *Generalized linear mixed models: a practical guide for ecology and evolution.* Trends in Ecology & Evolution, Volume 24, Issues 3, 127-135, 29 January 2009

Bosanquet N., Ferry J., Lees C., Thornton J. (2005) REFORM: *Maternity Services in the NHS.* Reform: London; 2005 www.reform.co.uk/filestore/pdf/Maternity%20services%20in%20the%20NHS.pdf

Bowes A. and Domokos T. (2003) Your dignity is hung up at the door: Pakistani and White women's experiences of childbirth. In: Earle S, Letherby G, editors. Gender, Identity and Reproduction: Social Perspectives. Basingstoke: Palgrave Macmillan; 2003. pp. 87–102.

Brace V., Penney G., Hall M. (2004) *Quantifying severe maternal morbidity: a Scottish population study*. BJOG: an International Journal of Obstetrics and Gynaecology. May 2004, Vol.111, pp.481-484

Bragg F., Cromwell D., Edozien L., Gurol-Urganci I., Mahmood T., et al. (2010) *Variations in rates of caesarean section among English NHS trusts after accounting for maternal and clinical risk: cross sectional study.* BMJ 2010; 341

Breslow NE. and Clayton DG. (1993) *Approximate inference in generalized linear mixed models*, J. Am. Statist. Assoc. 88, 9-25.

British Columbia Reproductive Care Program (2001) *Obstetric Guideline 14: Assisted vaginal birth: the use of forceps or vacuum extractor.* Vancouver, Canada: BCRCP; 2001. [http://www.csh.org.tw/Dr.TCJ/Educartion/Guideline/OB%20guideline/Assist.Delivery %20Guideline.pdf].

Brocklehurst P., Hardy P., Hollowell J., Linsell L., Macfarlane A., McCourt C, et al. (2011) *Perinatal and maternal outcomes by planned place of birth for healthy women with low risk pregnancies: the Birthplace in England national prospective cohort study.* BMJ 2011;343:d7400.

Brown SA. and Grimes DE. (1995) *A Meta-Analysis of Nurse Practitioners and Nurse Midwives in Primary Care.* Nursing Research, Nov/Dec 1995. Vol. 44, No 6.

Brown HC., Paranjothy S., Dowswell T., Thomas J. (2008) Package of care for active management in labour for reducing caesarean section rates in low risk women. Cochrane Database Syst Rev (4):CD004907

Brown S. (2012) Measures of Shape: Skewness and Kurtosis. Oak Road Systems. Available from http://www.tc3.edu/instruct/sbrown/stat/shape.htm#Skewness

Browne WJ. and Rasbash J. *Multilevel Modelling*. Institute of Education, University of London. http://www.cmm.bris.ac.uk/team/mmsage.pdf

Bryk AS. and Raudenbush SW. (1992) Hierarchical Linear Models: Application and Data Analysis Methods. SAGE: Newbury Park, CA

Burns EM., Rigby E., Mamidanna R., et al. (2012) *Systematic review of discharge coding accuracy.* J Public Health 2012; 34:138–48.

Caird J., Rees R., Kavanagh J., Sutcliffe K., Oliver K., Dickson K., Woodman J., Barnett-Page E., Thomas J. (2010) *The socioeconomic value of nursing and midwifery: a rapid systematic review of reviews.* EPPI Centre report 1801, London, EPPI Centre, Social Science Research Unit, Institute of Education, University of London.

Callaghan WM., Mackay AP., Berg CJ. (2008) *Identification of severe maternal morbidity during delivery hospitalizatuion, United States, 1991-2003.* Am J Obstet Gynecol 2008, 199(2):133

Campbell SE., Campbell MK., Grimshaw JM., Walker AE. (2001) *A systematic review of discharge coding accuracy.* J Public Health 2001; 23:205–11.

Cameron CA., Roberts CL., Bell J., Fischer W. (2007) *Getting an evidence-based post-partum heamorhhage policy into practice*. Aust N Z J Obstet Gynaecol 2007, 47(3):169-175.

Cantwell R., Clutton-Brock T., Cooper G., et al. (2011) *Saving Mother's Lives: reviewing maternal deaths to make motherhood safer 2006-2008*. BJOG 118 (Suppl 1):1-203

Carlin JB., Wolfe R., Brown CH., Gelman A. (2001) *A case study on the choice, interpretation and checking of multilevel models for longitudinal binary outcomes.* Biostatistics (2001),2 ,4 pp. 397-416.

Caton D., Corry MP., Frigoletto FD, Hopkins DP., Lieberman E., Mayberry L., et al. (2002) *The nature and management of labour pain: executive summary.* American Journal of Obstetrics and Gynecology 2002; 186(5):S1-S15.

CEMACH (2004) *Confidential Enquiry into Maternal and Child Health. Why Mothers Die 2000-2002*. RCOG Press, London.

CEMACH (2005) *Confidential Enquiry into Maternal and Child Health. Pregnancy in women with type 1 and type 2 diabetes in 2002-03, England, Wales and Northern Ireland.* CEMACH: London; 2005.

CEMACH (2007) *Saving Mothers' Lives: Reviewing maternal deaths to make motherhood safer 2003-2005. Summary and key recommendations for midwives.*

CEMACH (2008) *Perinatal Mortality 2006: England, Wales and Northern Ireland*. CEMACH, London, 2007.

CEMACH (2009) Perinatal Mortality 2007. London: CEMACH, 2009

Centre for Maternal and Child Enquiries (CMACE, 2010) *Maternal obesity in the UK: Findings from a national project*. London: CMACE 2010.

Centre for Maternal and Child Enquiries & Royal College of Obstetricians and Gynaecologists (2010) CMACE & RCOG Joint Guideline. Management of women with obesity in pregnancy. London: CMACE & RCOG, 2010.

Centre for Maternal and Child Enquiries (CMACE, 2011) *Saving Mothers' Lives: reviewing maternal deaths to make motherhood safer: 2006–08.* The Eighth Report on Confidential Enquiries into Maternal Deaths in the United Kingdom. BJOG 2011; 118 (Suppl. 1):1–203. Available from: http://onlinelibrary.wiley.com/doi/10.1111/j.1471-0528.2010.02847.x/pdf

Centre for Multilevel Modelling (CMM), Bristol, Online Workshops. http://www.bristol.ac.uk/cmm/software/support/workshops/materials/multi-br-print.pdf

Centre for Workforce Intelligence (CfWI, 2011) Workforce risks and opportunities: Nursing and Midwifery.

Chalmers I., Hetherington J., Newdick M., Mutch L., Grant A., Enkin M., Enkin E., Dickersin K. (1986) *The Oxford Database of Perinatal Trials: Developing a register of published reports of controlled trials.* Control Clin Trials 1986; 7:306–324.

Chalmers I., Enkin M., Keirse MJNC. eds. (1989) Effective Care in Pregnancy and Childbirth. Oxford: Oxford University Press, 1989.

Chalmers B., Kaczorowski J., Darling E., Heaman M., Fell DB., O'Brien B., et al. (2010) *Cesarean and vaginal birth in canadian women: a comparison of experiences*. Birth 2010;37:44–9.

Chamberlain G. (1991) Vital statistics of birth. Br Med J 1991 303: 178- 181

Chamberlain G., Wraight A., Crowley P. editors. (1994) Home births: the report of the 1994 confidential enquiry by the National Birthday Trust Fund. Carnforth, Lancs: Parthenon Publishing; 1997.

Chambliss LR., Daly C., Medearis AL., Ames M., Kayne M., Paul R. (1992) *The role of selection bias in comparing caesarean birth rates between physician and midwifery management*. Obstet Gynecol. 1992; 80:161-165.

Cho Sung-Hyun (2003) *Using Multilevel Analysis in Patient and Organizational Outcomes Research.* Nursing Research, Vol 52, No1., January/February 2003

Clinical Negligence Scheme for Trusts (2012) Maternity (clinical risk management standards, Ver 1, 2012/13).

Cluett ER., Nikodem VC., McCandlish R., Burns E. (2004) *Immersion in water in pregnancy, labour and birth (Cochrane Review).* In: The Cochrane library, Issue 2, 2004.

Cochrane A.L. (1979) 1931-1971: A critical review with particular reference to the medical profession. In: Teeling-Smith G, Wells N, eds. Medicines for the Year 2000. London: Office of the Health Economics, 1979.

Commission for Healthcare Audit and Inspection (HCC, 2008) Towards better births: A review of maternity services in England. ISBN 978-1-84562-194-0. Available from:http://image.guardian.co.uk/sys-files/Society/documents/2008/07/10/Towards_better_births.pdf

Cragin L. and Kennedy HP. (2006) *Linking obstetric and midwifery practice with optimal outcomes*. JOGNN, 35, 779-785; 2006

Darzi Lord (2008) *High quality care for all: NHS next stage review final report*. London: The Stationery Office; 2008.

Dattani N., Datta-Nemdharry P., Macfarlane A. (2011) *Linking maternity data for England, 2005–06: methods and data quality*. Health Stat Q 2011;49:53–79.doi:10.1057/hsq.2011.3.

Davis A. (2013) *Choice, policy and practice in maternity care since 1948. History and Policy.* Available from: http://www.historyandpolicy.org/papers/policy-paper-146.html .

Donabedian A. (1966) *Evaluating the quality of medical care.* Milbank Memorial Fund Quarterly, 44(1), 166-203.

Dowd JB. (2007) *Early childhood origins of the income/health gradient: the role of maternal health behaviors*. Soc Sci Med 2007;65:1202–13.

Department for Communities and Local Government DCLG (2011) The English Indices of Deprivation 2010. Neighbourhoods Statistical Release. https://www.gov.uk/government/publications/english-indices-of-deprivation-2010.

Department of Health (1993) Changing Childbirth: report of the Expert Maternity Group: HMSO, 1993.

Department of Health (2004) National Service Framework for Children, Young People and Maternity Services. No. 40498. Maternity Services, Standard 11. London. www.dh.gov.uk.

Department of Health (2005) Evidence to inform the National Service Framework for Children, Young People and Maternity Services, 2005, London.

Standing Maternity and Midwifery Advisory Committee (1970) Domiciliary Midwifery and Maternity Beds Needed. London: HMSO.

Department of Health (2007) Maternity Matters: Choice, Access and Continuity of Care in a Safe Service. No.278867. London April 2007. www.dh.gov.uk/publications.

Department of Health (2008) High quality care for all: NHS Next Stage Review final report. Professor the Lord Darzi of Denham KBE. Available from:

http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_085825.

Department of Health (2009) Transforming Community Services.

Department of Health (2009a) The Year: NHS Chief Executive's annual report 2008/09. London: Department of Health; 2009.

Department of Health (2009b) Delivering high quality midwifery care: the priorities, opportunities and challenges for midwives. London: DH; 2009. Available from: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_106063.

Department of Health (2009c) The operating framework for the NHS in England 2010/11. London: Department of Health; 2009.

Department of Health (2010a) Equity and Excellence: Liberating the NHS. Presented to Parliament by the Secretary of State for Health by Command of Her Majesty.

Department of Health (2010b) The NHS Outcomes Framework 2011/12. London: Department of Health; 2010.

Department of Health (2012a) The Mandate: A mandate from the Government to the NHS Commissioning Board: April 2013 to March 2015. London: Department of Health; 2012.

Department of Health (2012b) 2012/3 Choice Framework. London: Department of Health; 2012.

Department of Health (2012c) The NHS Outcomes Framework 2013/14. London: Department of Health; 2012.

Department of Health (2012d) Maternity Services Pathway Payment System: A Simple Guide 2012/13. Gateway reference: 17184. Available from: http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_133896.pdf

Devane D., Begley CM., Clarke M., Walsh D., Sandall J., Ryan P., et al. (2010) *Socio-economic value of the midwife: A systematic review, meta-analysis, meta-synthesis and economic analysis of midwife-led models of care.* London.

Dezateux C., Hockley C., Johnson J., et al. (2006) *Millennium cohort study: birth registration and maternity. Hospital Episode Statistics Linkage.* Centre for Longitudinal Studies Report, 2006. Published by the Centre for Longitudinal Studies, Bedford Group for Lifecourse and Statistical Studies, Institute of Education, University of London. In collaboration with the National Perinatal Epidemiology Unit and the Institute for Child Health.

Diez R. (2001) *Investigating neighborhood and area effects on health. Commentary*. American Journal of Public Health. November 2001, Vol 91, No. 11.

Dodwell M. and Newburn M. (2010) NCT: *Normal birth as a measure of the quality of care. Evidence on safety, effectiveness and women's experiences.* Available from:

http://www.nct.org.uk/sites/default/files/related_documents/Normalbirthasameasureofthequalityofcare V3.pdf

Duke S. and Copp G. (1992) *Hidden Nursing*. Nursing Times, 1992 Apr 22-28; 88(17):40-2.

Editorial: British Births 1970 (1976) *Br Med J 1976;1:854*

Elliott J. and Shepherd P. (2006) *Cohort Profile: 1970 British Birth Cohort (BCS70).* Int. J. Epidemiol. (August 2006) 35 (4): 836-843. doi: 10.1093/ije/dyl174

Engle WA. (2006) *A recommendation for the definition of "late preterm (near-term) and the birth weight-gestational age classification system.* Seminars in Perinatology 30(1):2-7. Available at: http://dx.doi.org/10.1053/j.semperi.2006.01.007

Enkin M., Keirse M.J.N.C., Chalmers I. (1989) A Guide to Effective Care in Pregnancy and Childbirth. Oxford: Oxford University Press, 1989.

Enkin M. (1994) *Risk in pregnancy: the reality, the perception, and the concept*. Birth 21(3): 131-4.

Enkin M., Glouberman S., Groff P., Jadad A., Stern A. (2006) *Beyond Evidence: The Complexity of Maternity Care.* Birth 33(4): 265-9.

Essex H., Green J., Baston H., Pickett K. (2013) *Which women are at an increased risk of a caesarean section or an instrumental vaginal birth in the UK: an exploration within the Millennium Cohort Study.* BJOG 2013; DOI: 10.1111/1471-0528.12177.

Fantini MP., Stivanello E., Frammartino B., Barone AP., Fusco D., Dallolio L., Cacciari P. and Perucci CA. (2006) *Risk adjustment for inter-hospital comparison of primary cesarean section rates: need, validity and parsimony.* BMC Health Services Research 2006, 6:100

Field A. (2009) Discovering statistics using SPSS. Sage Publications Ltd 2009, third edition. Chapter 19 Multilevel linear models.

Fitzpatrick KE., Kurinczuk JJ., Alfirevic Z., Spark P., Brocklehurst P., Knight M. (2012a) *Uterine rupture by intended mode of delivery in the UK: a national case-control study.* PLoS medicine 2012;9(3):e1001184.

Fitzpatrick KE., Sellers S., Spark P., Kurinczuk JJ., Brocklehurst P., Knight M. (2012b) *Incidence and risk factors for placenta accreta/increta/percreta in the UK: a national case-control study*. PloS one 2012;7(12):e52893.

Flynn M. and McKeown M. (2009) *Nurse staffing levels revisited: a consideration of key issues in nurse staffing levels and skill mix research.* Journal of Nursing Management, 2009, 17, 759-766.

Ford JB., Roberts CL., Simpson JM., Vaughan J., Cameron CA. (2007) *Increased postpartum haemorrhage rates in Australia*. Int J Gynaecol Obstet 2007, 98(3):237-243.

Ford E., Ayers S. (2009) *Stressful events and support during birth: the effect on anxiety, mood and perceived control*. J Anxiety.Disord. 2009;23(2):260-8.

Forster Della A., et al. (2006) *Staffing in postnatal units: is it adequate for the provision of quality care? Staff perspectives from a state-wide review of postnatal care in Victoria, Australia*. BMC Health Services Research, 6, 83.

Freeman GK., Woloshynowych M., Baker R., Boulton M., Guthrie B., Car J., et al. (2006) *Continuity of Care 2006: What Have We Learned Since 2000 and What are Policy Imperatives Now? Report for the National Co-ordinating Centre for NHS Service Delivery and Organisation R & D (NCCSDO)*. London: NCCSDO, 2007.

Garrod D., Beale V., Miller A. (2011) (on behalf of the Centre for Maternal and Child Enquiries). *Midwifery Summary.* 2011 Centre for Maternal and Child Enquiries (CMACE), BJOG 118 (Suppl. 1), e1–e11. Available from: http://www.centreformidwiferyeducation.ie/wp-content/uploads/2011/03/CMACE-2011-Saving-Mothers-Lives-Midwifery-summary.pdf

Garout M., Tilney H., Tekkis P., Aylin P. (2008) *Comparison of administrative data with the Association of Coloproctology of Great Britain and Ireland (ACPGBI) colorectal cancer database*. Int J Colorectal Dis 2008;23:155–63.

Garson GD. (2013) Hierarchical Linear Modeling. Guide and Applications. Chapter I: Fundamentals of Hierarchical Linear (Multilevel) Modeling. SAGE Publications, Inc. www.sagepub.com/upm-data/47528_ch_1.pdf

Geller SE., Rosenberg D., Cox SM., et al. (2004) *The continuum of maternal morbidity and mortality: factors associated with severity.* Am J Obstet Gynecol. 2004; 191:939-944

Gerova V., Griffiths P., Jones S., Bick D. (2010) *The association between midwifery staffing and outcomes in maternity services in England: observational study using routinely collected data.* http://www.kcl.ac.uk/nursing/research/nnru/publications/Reports/NNRUMidwiferyReportDec10.pdf

Goldstein H. (1987) Multilevel Models in Educational and Social Research. London: Griffin.

Goldstein H. (1991) *Nonlinear Multilevel Models with an Application to Discrete Response Data*. Biometrika 78:45-51.

Goldstein H. (1995) Multilevel statistical models. Edward Arnold, London. In electronic form at: www.arnoldpublishers.com/support/goldstein.htm

Goldstein H., Browne W. and Rasbash J. (2000) *Extensions to the Intra-Unit Correlation Coefficient to Complex Generalised Linear Multilevel Models*. http://multilevel.ioe.ac.uk/team/materials/icccglmm.pdf

Gopalani S., Bennett K., Critchlow C. (2004*) Factors predictive of failed operative vaginal delivery.* American Journal of Obstetrics and Gynecology 191:896–902.

Green JM., Curtis P., Price H., Renfrew M. (1998) *Continuing to Care. The organisation of midwifery services in the UK: a structured review of the evidence.* Hale: Books for Midwives Press; 1998.

Green JM., Baston H., Easton S. et al. (2003) *Greater expectations? Inter-relationships between women's expectations and experiences of decision making, continuity, choice and control in labour, and psychological outcomes: summary report.* Leeds: Mother & Infant Research Unit; 2003. Available from: www.leeds.ac.uk/miru/

Green JM. and Baston HA. (2007) *Have women become more willing to accept obstetric interventions and does this relate to mode of birth? Data from a prospective study.* Birth 2007;34:6–13.

Griffiths P., et al. (2008) *State of the art metrics for nursing: a rapid appraisal*. NNRU, King's College London

Griffiths P., Maben J. and Murrels T. (2011) *Organisational quality, nurse staffing and the quality of chronic disease management in primary care: observational study using routinely collected data*. Int J Nurs Stud. 2011 Oct;48(10):1199-210. doi: 10.1016/j.ijnurstu.2011.03.011. Epub 2011 May 14.

Guo G., and Zhao H. (2000) *Multilevel Modeling for Binary Data.* Annu. Rev. Sociol. 2000.26:441-62.

Gupta JK., Hofmeyr GJ., Shehmar M. (2012) *Position in the second stage of labour for women without epidural anaesthesia* . Cochrane Database of Systematic Reviews (5): CD002006.

Guyatt G. (1991) Evidence-based medicine. ACP J Club 1991;114-A-16.

Gyte G., Dodwell M., Newburn M., Sandall J., Macfarlane A., Bewley S. (2010) *Safety of planned home births. Findings of meta-analysis cannot be relied on.* BMJ 2010; 341:c4033.

de Jonge A., van der Goes By, Ravelli AC., Amelink-Verburg MP., Mol BW., Nijhuis JG., et al. (2009) *Perinatal mortality and morbidity in a national cohort of 529,688 low-risk planned home and hospital births.* BJOG, 2009, Aug; 116(9):1177-84.

Hall ES., Poynton MR., Narus SP., Jones SS., Evans RS., Varner MW., Thornton SN. (2009) *Patient - level analysis of outcomes using structured labor and delivery data.* J Biomed Inform (2009), doi 10.1016/j.jbi.2009.01.008

Hallowell J. (2011) Birthplace programme overview: background, component studies and summary of findings. Birthplace in England research programme. Final report. Part 1. NIHR Service Delivery and Organisation programme; 2011.

Hartmann K., Viswanathan M., Palmieri R., Gartlehner G., Thorp JJr., Lohr KN. (2005) *Outcomes of routine episiotomy: a systematic review*. JAMA.2005; 293:2141-2148

Hastings-Tolsma M., Vincent D., Park JH., Pan D. (2009) *Treatment Patterns and Outcomes in a Low-Risk Nurse-Midwifery Practice.* Applied Nursing Research 22 (2009) 10-17.

Hatem M., Sandall J., Devane D., Soltani H., Gates S. (2008) *Midwife-led versus other models of care for childbearing women*. Cochrane Database Syst Rev. 2008 Oct 8;(4):CD004667.

Heslehurst N, Ells LJ, Simpson H, Batterham A, Wilkinson J, Summerbell CD. (2007) *Trends in maternal obesity incidence rates, demographic predictors, and health inequalities in 36,821 women over a 15-year period.* BJOG: *An International Journal of Obstetrics and Gynaecology* 2007;114(2):187-94.

The Health and Social Care Information Centre, HSCIC, *Secondary Uses Service (SUS)*. http://www.hscic.gov.uk/sus

Health and Social Care Information Centre (HSCIC, 2011) Hospital Episode Statistics: NHS Maternity Statistics, 2010-11. NHS Maternity Statistics 2010-11 Explanatory notes. December 2011. https://catalogue.ic.nhs.uk/publications/hospital/maternity/nhs-mater-eng-2010-2011/nhs-mate-eng-2010-2011-rep.pdf

Health and Social Care Information Centre (HSCIC, 2012) Hospital Episode Statistics, Admitted Patient Care - England 2011-2012 [Internet]. Health and Social Care Information Centre; 2012.

Health and Social Care Information Centre (2013) NHS Maternity Statistics, England: 2012-13 Available at: http://www.hscic.gov.uk/catalogue/PUB12744.

Health and Social Care Information Centre (HSCIC, 2014) NHS Workforce Statistics in England, Non-medical staff – 2003-2013, As at 30 September: Non-medical bulletin tables [.xlsx]. March 25, 2014.

Heck RH., et al. (2012) *Multilevel Modeling of Categorical Outcomes Using IBM SPSS. Quantitative Methodology Series.* Routledge 2012.

Hilder L., Moser K., Dattani N et al. (2007) *Pilot linkage of NHS numbers for babies data with birth registrations.* Health Stat Q 2007;33:25–33

Hodnett ED. (1997) *Are nurses effective providers of labour support? Should they be? Can they be?* Birth 1997; 24:78-80.

Hodnett ED. (2002) *Home-like versus conventional institutional settings for birth* (Cochrane Review): In: The Cochrane Library

Hodnett ED., Gates S., Hofmeyr GJ., Sakala C. (2007) *Continuous support for women during childbirth. Cochrane Database of Systematic Reviews* 2007, Issue 3, Art.No.:CD003766. DOI: 10.1002/14651858. CD003766.pub2.

Hodnett ED., Stremler R., William AR., Weston JA., et al. (2008*) Effects on birth outcomes of a formalized approach to care in hospital labour assessment units: international, randomized controlled trial.* BMJ. 2008; 337: a1021

Hodnett ED., Downe S., Walsh D., Weston J. (2010) *Alternative versus conventional institutional settings for birth.* Cochrane Database Syst Rev. 2010; 9:CD000012.

Hodnett ED., Gates S., Hofmeyr GJ., Sakala C. (2013) *Continuous support for women during childbirth.* Cochrane Database of Systematic Reviews. The Cochrane Library 2013, Issue 7. http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD003766.pub5/pdf

Holt P., Poloniecki J., Thompson M. (2012) *Multicentre study of the quality of a large administrative data set and implications for comparing death rates.* Br J Surg 2012;99:58–65

House of Commons Health Committee (1992) *Maternity Services. Vol 1: Second report of session 1991-92 (Winterton Report)*. HMSO, London, 1992.

House of Commons Health Committee (2003) *Choice in Maternity Services. London.*

House of Commons Health Committee (2007) *Workforce Planning: Fourth report of session 2006-07.* London: TSO

House of Commons Health Committee (2008) *Modernising Medical Careers.* Third Report of Session 2007–08.

House of Commons Public Accounts Committee (2014) *Maternity services in England*.

Hox JJ. (1995) Applied multilevel analysis. Amsterdam:TT-Publikaties.

IBM SPSS Statistics (2010) *Make Smarter Decisions with Your Nested Data. Using GLMM with continuous and categorical targets.* IBM Software Business Analitics, 2010. http://public.dhe.ibm.com/common/ssi/ecm/en/ytw03081usen/YTW03081USEN.PDF

Iezzoni LI. (1997a) *Assessing quality using administrative data. Annals of Internal Medicine*, 127, 666-673

Iezzoni LI.(1997) Risk Adjustment for measuring healthcare outcomes. 2. Health Administration Press; 1997.

Iezzoni LI. (2003) Risk adjustment for measuring health care outcomes (Third ed.) Chicago: Health Administration Press, 2003

Institute of Medicine (2001) *Crossing the Quality Chasm.* Washington: National Academy Press

Janssen P., Lee S., Ryan E., Etches D., Farquharson D., Peacock D., Klein M. (2002) *Outcomes of planned home births versus planned hospital births after regulation of midwifery in British Columbia.* CMAJ, Feb.5, 2002; 166(3):315-23.

Janssen PA., Saxell L., Page LA., Klein MC., Liston RM., Lee SK. (2009) *Outcomes of planned home birth with registered midwife versus planned hospital birth with midwife or physician.* CMAJ. 2009 Sep 15; 181(6-7):377-83.

Jarman B., Gault S., Alves B, et al. (1999) *Explaining differences in English hospital death rates using routinely collected data.* BMJ. 1999;318:1515–1520.

Johal A., Cromwell D. and van der Meulen J. (2013) *Hospital Episode Statistics and Revalidation. Creating the Evidence to Support Revalidation.* Report by the Clinical Effectiveness Unit, Royal College of Surgeons of England (RCS) and London School of Hygiene and Tropical Medicine. http://www.rcseng.ac.uk/surgeons/research/surgical-research/docs/hospital-episode-statistics-and-revalidation-report-2013

Johanson R., Wilkinson P., Bastible A., et al. (1993) *Health after childbirth: a comparison of normal and assisted vaginal delivery.* Midwifery 1993;9(3):161-9.

Johanson RB. and Menon V. (1999) *Vacuum extraction versus forceps for assisted vaginal delivery.* Cochrane Database of Systematic Reviews, 1999, Issue 2. Article No: CD000224

Joseph KS., Rouleau J., Kramer MS., Young DC., Liston RM., Baskett TF. (2007) *Investigation of an increase in postpartum haemorrhage in Canada*. BJOG 2007, 114(6):751-759.

Joyce R., Webb R., and Peacock J. (2002) *Predictors of obstetric intervention rates: case-mix, staffing levels and organisational factors of hospital of birth.* J Obstet Gynaecol, 22(6), 618-25.

Kabiru W. and Raynor BD. (2004) *Obstetric outcomes associated with increase in BMI category during pregnancy*. American Journal of Obstetrics and Gynecology 191:928–932.

Kachman SD. (2000) An Introduction to Generalized Linear Mixed Models. Department of Biometry, University of Nebraska-Lincoln. http://statistics.unl.edu/faculty/steve/glmm/paper.pdf

Kane R. (1997) Understanding health care outcomes research. Gaithersburg, MD: Aspen Publishers, Inc.

Kane R., Shamlian T., Mueller C., Duval S., Wilt T. (2007) *The association of registered nurse staffing levels and patient outcomes: systematic review and meta-analysis*. Medical Care 45 (12), 1195-1294.

Keriakos R., Sugumar S., Hilal N. (2013) *Instrumental vaginal delivery – back to basics*. Journal of Obstetrics and Gynaecology, 2013; 33:781-786.

Balanda KP. and MacGillivray HL. (1988) *Kurtosis: A Critical Review.* The American Statistician 42:2 [May 1988], pp 111–119

Khan HR. and Ewart JH. (2011) *Multilevel Logistic Regression Analysis Applied to Binary Contraceptive Prevalence Data***.** Journal of Data Science **9**(2011), 93-110.

Knight M. (2007) *Eclampsia in the United Kingdom 2005*. BJOG 2007;114:1072-8.

Knight M., Kurinczuk JJ., Spark P. and Brocklehurst P. (2007) *United Kingdom Obstetric Surveillance System UKOSS Annual Report 2007.* National Perinatal Epidemiology Unit, Oxford.

Knight M., Kurinczuk JJ., Spark P., Brocklehurst P., on behalf of the UKOSS Steering Committee (2008) *Cesarean Delivery and Peripartum Hysterectomy*. Obstetrics & Gynecology 2008;111(1):97-105

Knight M., Callaghan WM., Berg C., Alexander S., et al. (2009a) *Trends in postpartum hemorrhage in high resource countries: a review and recommendations from the International Postpartum Hemorrhage Collaborative Group*. BMC Pregnancy and Childbirth 2009, 9:55

Knight M., Kurinczuk JJ., Spark P. and Brocklehurst P. (2009b) *Inequalities in maternal health: national cohort study of ethnic variation in severe maternal morbitities*. BMJ 2009;338:b542.

Knight HK., Gurol-Urganci I., Mahmood TA., Richmond D., Templeton A., van der Meulen JH., Cromwell D. (2013) *Evaluating maternity care using national routine datasets: How are statistics affected by the quality of data on method of delivery?* BMC health services research, 13 (1). p. 200. ISSN 1472-6963 DOI: 10.1186/1472-6963-13-200

Kreft I., de Leeuw J., Aiken L. (1995) *The effect of different forms of centering in hierarchical linear modeling*. Multivariate Behavioral Research. 1995;30:1–20

Kreft I. and De Leeuw J. (1998) Introducing multilevel modelling. London: Sage Publications.

King's Fund (2008) *Safer Births: Everybody's Business*. London 2008

King's Fund (2011) *Staffing in Maternity Units. Getting the right people in the right place at the right time*. London 2011.

Kirsop R., Jenkins GJ., Saunders DM., Houghton CR. (1992) *The influence of maternal age on cesarean section rates.* Aust N Z J Public Health 1992;32(3):206– 207.

Lalonde A., Daviss BA., Acosta A., Herschderfer K. (2006) *Postpartum haemorrhage today: ICM/FIGO initiative 2004-2006*. Int J Gynaecol Obstet 2006, 94(3):243-253

Lankshear AJ., Sheldon TA., Maynard A. (2005) *Nurse staffing and healthcare outcomes: a systematic review of the international research evidence.* ANS Adv Nurs Sci Apr-Jun 2005;28(2):163-74.

Lewis G. (2007, Ed.) *The Confidential Enquiry into Maternal and Child Health (CEMACH). Saving Mothers' Lives: Reviewing Maternal Deaths to Make Motherhood Safer 2003-2005.* The Seventh Report on Confidential enquiries into Maternal Deaths in the UK, CEMACH, London.

Lindgren HE., Radestad IJ., Christensson K., Hildingsson IM. (2008) *Outcome of planned home births compared to hospital births in Sweden between 1992 and 2004. A population-based register study.* Acta Obstet Gynecol Scand. 2008; 87(7):751-9.

Lohr KN. (1990, Ed.) *Medicare: A strategy for quality assurance*. Vols I and II. Washington, DC:National Academy Press, 1990.

London LSA Annual report to NMC (2011) available from:http://www.nmc-uk.org/Documents/Midwifery-LSA-reports/LSAreports2010-2011/LSA_London_2010-2011.pdf

Loudon JAZ., Groom KM., Hinkson L., Harrington D., Pareson-Brown S. (2010) *Changing trends in operative delivery performed at full dilation over a 10-year period.* JOG, May 2010; 30(4):370-375.

MacCullagh P. and Nelder JA. (1989) Generalized Linear Models. Chapman and Hall, 1989.

MacDorman MF. and Singh GK. (1998) *Midwifery care, social and medical risk factors, and birth outcomes in the USA.* Journal of Epidemiology & Community Health, 52. 310-317.

Macfarlane A. and Mugford M. (2000) Birth Counts: Statistics of Pregnancy and Childbirth. London: The Stationery Office, 2000.

Majoko F. and Gardener G. (2012) *Trial of instrumental delivery in theatre versus immediate caesarean section for anticipated difficult assisted births* (Review). *The Cochrane Library* 2012, Issue 10.

Maloney K. (1999) *An overview of outcomes research and measurement*. Journal for healthcare quality; 21 (6): 4-10

Mark BA., Harless DW., McCue M., Xu Y. (2004) *A longitudinal examination of hospital registered nurse staffing and quality of care*. HSR. 2004;39:279–301.

Main EK., Moore D., Farrell B., Schimmel LD., Altman RJ., Abrahams C., et al. (2006) *Is there a useful caesarean birth measure? Assessment of the nulliparous term singleton vertex caesarean birth rate as a tool for obstetric quality improvement*. Am J Obstet Gynecol 194:1644-51

Mainz J. (2003) *Defining and classifying clinical indicators for quality improvement.* International Journal for Quality in Health Care, Vol.15, N6:pp523-530

Mainz J. (2004) *Quality Indicators: essential for quality improvement*. Editorial. International Journal for Quality in Health Care 2004; Volume 16, Supplement 1: pp. i1–i2. International Society for Quality in Health Care and Oxford University Press 2004

Mason WM., Wong GM. and Entwistle B. (1983) *Contextual Analysis through the Multilevel Linear Model.* Sociol. Methodol. 13:72-103.

Maternity Services Committee (1959) The Cranbrook Report.

Maternity and Midwifery Advisory Committee (1967) The Peel Report.

Maternity Care Working Party (2007) Making normal birth a reality: Consensus statement from the Maternity Care Working Party: NCT/RCM/RCOG;2007

McGillis H., Doran D., Pink G. (2004) *Nurse staffing models, nursing hours, and patient safety outcomes*. J Nurs Admin. 2004;34:41–45.

McMahon JM., Pouget ER., Tortu S. (2006) *A guide for multilevel modeling of dyadic data with binary outcomes using SAS PROC NLMIXED.* Comput Stat Data Anal 50: 3663–3680. [PMC free article]

Mead MM. and Kornbrot D. (2004) *The influence of maternity units' intrapartum intervention rates and midwives' risk perception for women suitable for midwifery-led care*. Midwifery. 2004 Mar; 20(1): 61-71.

Mead M., Bogaerts A., Reyns M., Poutas M., Hel S. (2006) *Midwives' perception of intrapartum risk in England, Belgium and France.* Eur Clinics Obstet Gynaecol (2006) 2:91-98.

Medicine I of. (2001) *Crossing the quality chasm: A new health system for the 21st century*. 1st ed. Washington: National Academy Press; 2001.

MIDIRS and the NHS Centre for Reviews and Dissemination (1999) *Support in labour: informed choice for professionals leaflet.* Midwives Information and Resource Service. Bristol, England, 1999.

Midwifery 2020 (2010a) *Measuring Quality Workstream Final Report*. Midwifery 2020 Programme. Available from: http://www.midwifery2020.org/documents/2020/Measuring_Quality.pdf

Midwifery 2020 (2010b) Delivering Expectations. London: Department of Health; 2010.

Morrell CJ., Spiby H., Stewart P., Walters S., Morgan A (2000) *Costs and effectiveness of community postnatal support workers: randomised controlled trial*. BMJ, vol 321, pp 593-8

Mori R., Dougherty M., Whittle M. (2008) *An estimation of intrapartum-related perinatal mortality rates for booked home births in England and Wales between 1994 and 2003.* BJOG. 2008 Apr; 115(5):554-9.

Murphy PA and Fullerton JT. (2001) *Measuring outcomes of midwifery care: Development of an instrument to assess optimality*. Journal of Midwifery and Women's Health, 46, 274-284.

Murphy DJ., Liebling RE., Verity L., Swingler R., Patel R. (2001a) *Early maternal and neonatal morbidity associated with operative delivery in second stage of labour: a cohort study*. Lancet 358:1203-1207.

Murphy DJ., Liebling RE., Patel R., Verity L., Swingler R. (2003) *Cohort study of operative delivery in the second stage of labour and standard of obstetric care*. *BJOG* 2003;110:610–5.

Murphy PA., Fullerton JT. (2006*) Development of the Optimality Index as a new approach to evaluating outcomes of maternity care.* JOGNN, 35, 770-778; 2006.

Murray J., Saxena S., Modi N., Majeed A., Aylin P., Bottle A. (2012) *Quality of routine hospital birth records and the feasibility of their use for creating birth cohorts*. J Public Health (Oxf). 2012;35:298–307.

National Audit Office (2013) Maternity Services in England. Report by the Comptroller and Audit General, DoH 2013, London: the Stationary Office

National Childbirth Trust (2002) The NCT birth policy: 12 principles. London:NCT

National Institute for Health and Clinical Excellence (NICE) (2004). *Caesarean Section.* National Collaborating Centre for Women's and Children's Health, RCOG Press; 2004.

National Institute for Health and Clinical Excellence (NICE) (2007). *Intrapartum care of healthy women and their babies during childbirth.* London. National Collaboration Centre for Women's and Children's Health, RCOG, 2007.

NICE Clinical Guideline 70 (2008) *Induction of Labour.* National Collaborating Centre for Women's and Children's Health. Commissioned by National Institute for Health and Clinical Excellence, July 2008. Published by the RCOG Press.http://www.nice.org.uk/nicemedia/pdf/CG070NICEGuideline.pdf. http://www.nice.org.uk/nicemedia/live/12012/41255/41255.pdf

National Institute for Health and Clinical Excellence (NICE) (2011) *Caesarean Section.* CG132. London: National Collaborating Centre for Women's and Children's Health.

Needleman J. (2001) *Nurse staffing and patient outcomes in hospitals.* Final Report for Health Resources Services Administration 2001; Contract No. 230990021.

Needleman J., Buerhaus P., Mattke S., Stewart M., Zelevinsky K. (2002) *Nurse-staffing levels and the quality of care in hospitals*. NEJM. 2002;346:1715–1722.

NHS Institute for Innovation and Improvement (2006).. *Delivering quality and value. Pathways to success: a self-improvement toolkit. Focus on normal birth and reducing Caesarean section rates.* Coventry: NHS Institute.

National Patient Safety Agency (2006).

National Health Service (2006) NHS Maternity Statistics, England: 2004–05. London: National Health Service; 2006.

NHS Employers (2009) Agenda for Change at a glance. London, NHS Employers (www.nhsemployers.org).

NHS Litigation Authority (2006) Factsheet 5: 2005–06. Available from: www.nhsla.com

NHS Litigation Authority (2012) Ten years of maternity claims. An analysis of NHS Litigation Authority data. NHSLA London 2012.

NHS Information Centre, Workforce and Facilities Team (2011). NHS Workforce: Summary of staff in the NHS: Results from September 2010 Census. March 2011.https://catalogue.ic.nhs.uk/publications/workforce/numbers/nhs-staf-non-medi-2000-2010/nhs-staf-non-medi-cens-bull-2000-2010-rep.pdf

NHS Information Centre (2012) NHS maternity statistics, England: 2011-2012. Department of Health.

NHS Information Centre (2012a) NHS Workforce: Summary of staff in the NHS: Results from September 2011 Census. Available from:

http://www.hscic.gov.uk/catalogue/PUB05221/nhs-staf-2001-2011-medi-dent-work-rep.pdf

NHS Information Centre (2013a) Indicators for Quality Improvement [Internet]. 2013. Available from: https://mqi.ic.nhs.uk/

NHS Workforce Review Team (2009) *Assessment of Workforce Priorities 2009/10.*

North Staffordshire Changing Childbirth Research Team (2000) *A randomised study of midwifery caseload care and traditional 'shared care'*. Midwifery 2000;16(4):295–302.

Nursing and Midwifery Council (2004) Midwives Rules and Standards. London: NMC; 2004. www.nmc-uk.org/aFrameDisplay.aspx?DocumentID=169.

Nursing and Midwifery Council (2006) Midwives and home birth. NMC Circular 8-2006. London: NMC; 2006 www.nmc-uk.org/aFrameDisplay.aspx?DocumentID=1472.

O'Connell R. and Downe S. (2009) *A metasynthesis of midwives' experience of hospital practice in publicly funded settings: compliance, resistance and authenticity.* Health (London) 2009; 13; 589. Published by Sage.

Office for National Statistics (2009) *Statistical Bulletin: Who is having babies?*

Office for National Statistics (2011) Public Service Output, Inputs and Productivity: Healthcare 2011. Available from: http://www.ons.gov.uk/ons/rel/psa/public-service-productivity/healthcare-2011/public-service-output--input-and-productivity.pdf

Office for National Statistics (2012) Conceptions in England and Wales, 2010.

Office for National Statistics (2013a) Death Registrations Summary Tables, England and Wales, 2012, July 2013.

Office for National Statistics (2013b) Child Mortality Statistics: Childhood, Infant and Perinatal, 2011, February 2013.

Office for National Statistics (2013c) Statistical bulletin: Live Births in England and Wales by Characteristics of Mother 1, 2012.

Office for National Statistics (2013d) Statistical bulletin: Births in England and Wales by Parents' Country of Birth, 2012

Office for National Statistics (2013e) Birth Summary Tables, England and Wales, 2012.

Olde E., van der Hart O., Kleber R., van Son M. (2006) *Posttraumatic stress following childbirth: review.* Clin Psychol Rev 2006;26:1–16.

Olsen O. (1997) *Meta-analysis of the safety of home births.* Birth. 1997 Mar; 24(1):4-13.

Olsen O., Jewell MD. (1998) *Home versus hospital birth.* Cochrane Database of Systematic Reviews. 1998 (3):CD000352.

O' Mahony F., Hofmeyr GJ., Menon V. (2010) *Choice of instruments for assisted vaginal delivery*. Cochrane Database of Systematic Reviews (11): CD005455 .

Page L., McCourt C., Blake S., Beake S., Vail A., Hewison J. (1999) *Clinical interventions and outcomes of One-to-One midwifery practice.* Journal of Public Health Medicine, 1999, vol.21, No 3, pp.243-248

Page L., Beake S., Vail A., et al. (2001) *Clinical outcomes of one-to-one midwifery practice.* Briitish Journal of Midwifery 2001; 9(11):700-6

Page L. (2003) *'When normal birth is in the minority'*, British Journal of Midwifery, 11 (6), 356-56.

Paranjothy S., Frost C., Thomas J. (2005) *How much variation in CS rates can be explained by case mix differences?* BJOG 2005: 112:658-66.

Parrish KM., Holt VL., Easterling TR., Connell FA., LoGerfo JP. (1994) *Effect of changes in maternal age, parity, and birth weight distribution on primary cesarean delivery rates [see comments]*. JAMA 1994; 271(6):443–447.

Patel RR. & Murphy DJ. (2004) *Forceps delivery in modern obstetric practice*. British Medical Journal 328: 1302– 1305.

Pattinson RC., Say L., Makin JD., et al. (2005) *Critical incident audit and feedback to improve perinatal and maternal mortality and morbidity. Cochrane Database of Systematic Reviews.* 2005:CD002961.

Peaceman AM., Feinglass J. and Manheim LM. (2002) *Risk-Adjustment of Caesarean Delivery Rates: A Practical Method for Use in Quality Improvement*. American Journal of Medical Quality. Vol 7, No3, May/June 2002.

Pencheon D. (2008) *The good indicators guide: how to use and choose indicators.* London: NHS Institute for Innovation and Improvement, 2008.

Penney G., Kernaghan D., Brace V. (2007) *Severe maternal morbidity – the Scottish experience 2003-2005. Saving Mothers' Lives: Reviewing maternal deaths to make motherhood safer 2003-2005*. London CEMACH 2007. http://www.cmace.org.uk/Publications/CEMACH-Publications/Maternal-and-Perinatal-Health.aspx

Petrou S., Kupek E., Vause S., Maresh M. (2001) *Clinical, provider and sociodemographic determinants of the number of antenatal visits in England and Wales*. Soc Sci Med 2001;56:1123– 1134.

Petrou S. and Glazener C. (2002) *The economic costs of alternative modes of delivery during the first two months postpartum: results from a Scottish observational study.* BJOG 2002;109:214–7.

Powers DA. (2012) *Multilevel models for binary data.* New Directions for Institutional Research, Volume 2012, Issue 154, Summer 2012, Pages: 57–75, Article first published online : 19 JUN 2012, DOI: 10.1002/ir.20014

Prechtl HFR. (1968) *Neurological findings in newborn infants after pre- and para-natal complications*. In J. Jonxis, H.Visser, & J.A.Troelstra (Eds.), *Aspects of prematurity and dismaturity*. Leiden, The Netherlands: Stenfert Kroese

Prechtl HFR. (1980) *The optimality concept (Editorial).* Early Hum Dev, 1980; 4:201-5

Prendiville WJ., Elbourne D., McDonald S. (2000) *Active versus expectant management in the third stage of labour. Cochrane Database of Systematic Reviews 2000*, 3:Art No.: CD 000007

Prime Minister's Commission on the Future of Nursing and Midwifery in England (2010). Front Line Care: the future of nursing and midwifery in England. http://cnm.independent.gov.uk

Rabe-Hesketh S. and Skrondal A. (2008) Multilevel and Longitudinal Modeling Using STATA. A STATA Press Publications. STATACorp LP, College Station, Texas. Second Edition.

Rafferty A M., Clarke SP., Coles J., et al. (2007) *Outcomes of variation in hospital nurse staffing in English hospitals: cross-sectional analysis of survey data and discharge records*', Int J Nurs Stud, 44 (2).175-82.

Raisler J. (2000) *Midwifery care research: what questions are being asked? What lessons have been learned?* Journal of Midwifery and Women's Health. Vol.45, No 1, January/February 2000.

Raudenbush SW., and Bryk AS. (2002) Hierarchical Linear Models: Applications and Data Analysis Methods (2nd ed.) (Advanced Quantitative Techniques in the Social Sciences Series, No. 1). SAGE Publications; Thousand Oaks, CA.

Redshaw M. and Harris A. (1995) *Quality and quantity: staffing and skill mix in neonatal care*. Nurs Times. 1995 Jul 5-11; 91(27): 29-31

Redshaw M., Rowe R., Hockley C., Brocklehurst, P. (2006) *Recorded delivery: a national survey of women's experience of maternity care 2006.* National Perinatal Epidemiology Unit. Oxford; NPEU; 2007. ISBN 978 0 9735860 8 0. www.npeu.ac.uk

Redshaw M., Rowe R., Schroeder L., Puddicombe D., Macfarlane A., Newburn M., McCourt C., Sandall J., Silverton L., Marlow N. (2011) on behalf of the Birthplace in England Collaborative Group. Mapping maternity care: the configuration of maternity care in England Birthplace in England research programme. Final report part 3

Reinharz D., Blais R., Fraser W. (2000) *Cost-effectiveness of midwifery services versus medical services in Quebec.* Canadian Journal of Public Health 91(1): 112-5.

Richardson A. and Mmata C. (2007) NHS Maternity Statistics, England: 2005-06. The Information Centre; 2007. Available from: http://www.ic.nhs.uk/statistics-and-data-collections/hospital-care/maternity/nhs-maternity-statistics-2005-06

Roberts CL., Torvaldsen S., Cameron CA., Olive E. (2004) *Delayed versus early pushing in women with epidural analgesia: a systematic review and metaanalysis*. British Journal of Obstetrics and Gynaecology 111: 1333– 1340.

Robinson WS. (1950) *Ecological correlations and the behavior of individuals.* American Sociological Reviews, 15, 351-357.

Rodriguez G. and Goldman N. (2001) *Improved Estimation Procedures for Multilevel Models with Binary Response: A Case Study.* Journal of the Royal Statistical Society. Series A (Statistics in Society). Vol. 164, No. 2 (2001), pp. 339-355. Published by: Wiley Article Stable URL: http://www.jstor.org/stable/2680556

Rosenberg TJ., Garbers S., Lipkind H., Chiasson MA. (2005) *Maternal obesity and diabetes as risk factors for adverse pregnancy outcomes: differences among 4 racial/ethnic groups.* Am J Public Health 2005;95:1545-51.

Royal College of Midwives (2005) RCM Annual Survey of UK Heads of Midwifery Service. London: RCM; 2005.

Royal College of Midwives (2006) Position Paper 26 and Guidance Paper 26a. *Refocusing the Role of the Midwife*. London: RCM; 2006 www.rcm.org.uk/info/docs/PP%2026%20Refocusing%20the%20role%20of%20the%20midwife.doc

Royal College of Midwives (2010a) Evidence to the NHS Pay Review Body. 2010 Submission. Available at: www.rcm.org.uk/college/support-at-work/pay-pensions/pay/

Royal College of Midwives (2010b) Maternity Support workers: position statement. Royal College of Midwives website.

Royal College of Midwives (2011) State of Maternity Services Report 2011. London:RCM

Royal College of Nursing (2006) Setting appropriate ward nurse staffing levels in NHS acute trusts. London: Royal College of Nursing.

Royal College of Obstetricians and Gynaecologists (2001) The National Sentinel Caesarean Section Audit Report. London: RCOG Clinical Effectiveness Support Unit, 2001. [: ISBN 1–900364–66–2.]

Royal College of Obstetricians and Gynaecologists (2005) The Future Role of the Consultant. London: Royal College of Obstetricians and Gynaecologists

RCM, RCOG, NCT (2007) Making Normal Birth a Reality. Consensus statement from the Maternity Care Working Party. Available from: http://www.rcog.org.uk/files/rcog-corp/uploaded-files/JointStatmentNormalBirth2007.pdf

Royal College of Obstetricians and Gynaecologists (2007) Safer Childbirth: Minimum Standards for the Organisation and Delivery of Care in Labour. Royal Colleges of Midwives, Obstetricians and Gynaecologists, Anaesthetists, and Paediatrics and Child Health. Available from: http://www.rcog.org.uk/files/rcog-corp/uploaded-files/WPRSaferChildbirthReport2007.pdf

RCOG Hospital Recognition Committee (2007) Information from Dr TA Mahmood FRCOG, Chairmain of RCOG Hospital Recognition Committee.

Royal College of Obstetricians and Gynaecologists (2008) Standards for Maternity Care: report of a working party. Royal College of Obstetricians and Gynaecologists, Royal College of Midwives, Royal College of Anaesthetists et al. London: RCOG Press 2008. http://www.rcog.org.uk/files/rcog-corp/uploaded-files/WPRMaternityStandards2008.pdf

Royal College of Obstetricians and Gynaecologists (2009) Improving Patient Safety. Risk management for maternity. Clinical governence advice N2, 2009.

Royal College of Obstetricians and Gynaecologists (2009a) The Future Workforce in Obstetrics and Gynaecology: England and Wales full report. London: Royal College of Obstetrics and Gynaecologists.

Royal College of Obstetricians and Gynaecologists (2009b) Responsibility of Consultant On-call (Good Practice No8). London RCOG

Royal College of Obstetricians and Gynaecologists (2009/2011). Prevention and management of postpartum haemorrhage. Green-top guideline No 52 May 2009. Available from: http://www.rcog.org.uk/files/rcog-corp/GT52PostpartumHaemorrhage0411.pdf

Royal College of Obstetricians and Gynaecologists (2011) High Quality Women's Health Care : A proposal for change. London; 2011.

Royal College of Obstetricians and Gynaecologists (2011a) Operative vaginal delivery Green-top Guideline N26. January 2011

Royal College of Obstetricians and Gynaecologists (2013) Patterns of maternity care in English NHS hospitals 2011/12. London School of Hygiene & Tropical Medicine. Published by RCOG 2013.

Renfrew MJ., Hannah W., Albers L., Floyd E. (1998) *Practices that minimize trauma to the genital tract in childbirth: a systematic review of the literature*. Birth, 1998; 25:143-160.

Roberts LL., Ely JW., Ward MM. (2007) *Factors contributing to maternal birth-related trauma.* American Journal of Medical Quality 2007; 22; 334-343

Roberts CL., Cameron CA., Bell JC., Algert CS., Morris JM. (2008) *Measuring maternal morbidity in routinely collected health data. Development and validation of a maternal morbidity outcome indicator.* Medical Care 2008. Vol.46, N 8, pp 786-794.

Roberts CL., Ford JB., Algert CS., Bell JC., Simpson JM., Morris JM. (2009) *Trends in adverse maternal outcomes during childbirth: a population-based study of severe maternal morbidity.* BMC *Pregnancy and Childbirth*, February 2009, 9:7

Rodríguez G. and Goldman N. (2001) *Improved Estimation Procedures for Multilevel Models with Binary Response: A Case Study*. Journal of the Royal Statistical Society. Series A (Statistics in Society) Vol. 164, No. 2.

Ronsmans C. and Graham WJ. (2006) *Maternal Mortality: who, when, and why.* Lancet 2006 Sep 30;368 (9542):1189-200.

Rooks JP. (1999) *The Midwifery Model of Care*. Journal of Nurse-Midwifery. Vol. 44, No. 4, July/August 1999

Rosenblatt RA., Dobie SA., Hart G., Schneeweiss R., Gould D., Raine TR., Benedetti TJ., Pirani MJ. (1997) *Interspeciality Differences in the Obstetric Care of Low-Risk Women.* Am J Public Health, 1997; 87:344-351

Rosser J. and Anderson T. (2000) *What next? Taking normal birth out of the labour ward*. The Practising Midwife 3(4): 4-5.

Sakala C. (2006) *Vaginal or caesarean births: a systematic review to determine what is at stake for mothers and babies*. New York: Childbirth Connection. Available http://www.childbirthconnection.org/article.asp?ck=10271&ClickedLink=200&area=2

Sakala C. and Corry M. (2008) *Evidence-Based Maternity Care: What it is and what it can achieve*. Milbank Memorial Fund 2008.

Sandall J. (1997) *Midwives' burnout and continuity of care.* Br J Midwifery 1997; 5(2):106-111.

Sandall J. (1998) *Occupational burnout in midwives: new ways of working and the relationship between organizational factors and psychological health and well-being.* Risk Decision and Policy 3, 213-232.

Sandall J., Manthorpe J., Mansfield A., Spencer L. (2007) *Support Workers in Matternity Services: A national scoping study of NHS trusts providing maternity care in England 2006.* Final report. London: King's College.

Sandall J., Homer C., Sadler E., Rudisill C., Bourgeault I., Bewley S., et al. (2011) Staffing in maternity units. London: The King's Fund; 2011.

Sandall J., Soltani H., Gates S., Shennan A., Devane D. (2013) *Midwife-led continuity models versus other models of care for childbearing women*. Cochrane database Syst Rev [Internet]. 2013 Jan [cited 2014 Jan 20];8:CD004667. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23963739

Saunders D., Boulton M., Chapple J., Ratcliffe J., Levitan J. (2000) *Evaluation of Edgware Birthing Centre*. Barnett Health Authority: London.

Say L., Pattinson RC., Gulmezoglu AM. (2004) *WHO systematic review of maternal morbidity and mortality: the prevalence of severe acute maternal morbidity (near miss).* Reprod Health. 2004; 1:3.

Scott N., Williams G., Parker M. (2011) HESteria or hype? BMJ 2011; 343:d5947.

Sheridan M. and Sandall J. (2010) *Measuring the best outcome for the least intervention: can the Optimality Index-US be applied in the UK.* Midwifery (2010), doi:10.1016/j.midw.2009.09.002

Silber JH., Rosenbaum PR., Ross RN. (1995) Comparing the contributions of groups of predictors: Which outcomes vary with hospital rather than patient characteristics? Journal of the American Statistical Association. 1995;90(429):7–18

Silber JH. and Rosenbaum PR. (1997) A spurious correlation between hospital mortality and complication rates: the importance of severity adjustment. Med Care. 1997;35(10, Suppl.):OS77–OS92

Sinclair MK., Gardner J., Gillen P., Boreland Z., Hood R. (2001) *Analyses of the records of maternity patients deemed as low-risk on admission and of former patients' perceptions of maternity care.* Report submitted to Down Lisburn Trust, Northern Ireland.

Sinha S., Peach G., Poloniecki JD., Thompson MM. and Holt PJ. (2012) *Studies using English administrative data (Hospital Episode Statistics) to assess health-care outcomes – systematic review and recommendations for reporting.* The European Journal of Public Health, (1-6) Advance Access, published May 10, 2012. Published by Oxford University Press on behalf of EJPH.

Smith LJ. (2007) *Impact of birthing practices on the breastfeeding dyad*. J Midwifery Womens Health 2007;52:621–30.

Smith AHK., Dixon AL. and Page LA. (2009) *'Health-care professionals' views about safety in maternity services: a qualitative study*', *Midwifery,* 25 (1), 21-31.

Snijders TAB. and Bosker RJ. (1994) *Modeled variance in two-level models*. Sociol Methods Res. 1994;22:342-363

Snijders TAB. and Bosker RJ. (1999) Multilevel analysis: An introduction to basic and advanced multilevel modeling. London: Sage Publications.

Spencer C., Murphy DJ. and Bewley S. (2006) *Caesarean delivery in the second stage of labour: better training in instrumental delivery may reduce rates.* BMJ 2006;333(7569):613-4

Stivanello E., Rucci P., Carretta E., Oieri G. and Fantini PM. (2013) *Risk adjustment for cesarean delivery rates: how many variables do we need? An observational study using administrative databases.* BMC Health Serv Res. 2013; 13: 13. Published online 2013 January 10. doi: 10.1186/1472-6963-13-13, PMCID: PMC3554564.

Stout J. (2007) *Innovations Matter: Examples of support workers in maternity services*. London: CSIP.

Symon A., Winter C., Inkster M., Donnan P., (2009) *Outcomes for births booked under an independent midwife and births in NHS maternity units: matched comparison study.* BMJ 2009; 338:b2060

Temple J. (2010) *Time for Training: A Review of the Impact of the European Working Time Directive on the quality of training*. Available from: http://www.mee.nhs.uk/PDF/14274%20Bookmark%20Web%20Version.pdf

Thompson JE. (1986) *Nurse-midwifery care: 1925-1984*. Annu Rev Nurs Res 1986; 4:153-73

Thomas J., Paranjothy S., and Royal College of Obstetricians and Gynaecologists Clinical Effectiveness Support Unit (2001) *National sentinel caesarean section audit report.* London: RCOG Press; 2001.

Tourangeau AE., Giovannetti P., Tu JV., Wood M. (2002) *Nursing-related determinants of 30-day mortality for hospitalized patients*. Can J Nurs Res. 2002;33:71–88.

Tranmer M. and Elliot M. (2007) *Multilevel Modelling Coursebook*. Cathie Marsh Centre for Census and Survey Research, Manchester.

Unruh L. (2003) *Licensed nurse staffing and adverse events in hospitals*. Med Care. 2003;41:142–152.

Villar J., Valladares E., Wojdyla D., Zavaleta N., Shah A., Campodónico L., et al. (2006) *Caesarean delivery rates and pregnancy outcomes: the 2005 WHO global survey on maternal and perinatal health in Latin America.* Lancet 2006; 367 (9525): 1819–29.

Vincent C., Taylor-Adams S., Chapman EJ., Hewett D., Prior S., Strange P. et al. (2002) *How to investigate and analyse clinical incidents: clinical risk unit and association of litigation and risk management protocol .* British Medical Journal 320: 777 – 781.

Vincent C. (2006) Patient Safety. Edinburgh: Churchill Livingstone.

Walsh D. and Downe SM. (2004) *Outcomes of free-standing, midwife-led birth centres: a structured review.* Birth, 2004 Sep; 31(3):222-9.

Waterstone M., Bewley S. and Wolfe C. (2001) *Incidence and predictors of severe obstetric morbidity: case-control study.* BMJ 2001; 322:1089-94.

Wax JR., Lucas FL., Lamont M., Pinette MG., Cartin A., Blackstone J. (2010) *Maternal and newborn outcomes in planned home births vs planned hospital births: a meta-analysis.* AJOG. 2010 Sep; 203(3):243 e1-8.

Weir N. (2004) *Measuring and improving the quality of care*. Practical Neurology 2004;4;340-349

World Health Organisation (1985) Appropriate technology for birth. Lancet 1985; 2(8452):436-7

World Health Organization (1996) Care in Normal Birth: A Practical Guide. WHO Report of a Technical Working Group. Publication no.WHO/FRH/MSM/96.24. Geneva: Author, 1996.

World Health Organization Technical Working Group (1997) *Care in Normal Birth: A Practical Guide.* Short Communication. Birth, 1997 June; 24 (2):121-123.

World Health Organization & UNICEF (2004) Low Birthweight: Country,regional and global estimates. UNICEF, New York, 2004.

World Health Organization (2009) Monitoring Emergency Obstetric Care: a handbook. World Health Organization 2009. http://whqlibdoc.who.int/publications/2009/9789241547734_eng.pdf

World Health Organization (2011) Guidelines on optimal feeding of low birthweight infants in low-and middle-income countries.

Wiegers TA., Keirse MJNC., Berghs GAH., van der Zee J. (1996) *An approach to measuring quality of midwifery care.* J Clin Epidemiol 1996; 49:319-25

Wiegers TA., Keirse MJNC., Berghs GAH., van der Zee J. (1996a) Outcome of planned home and planned hospital births in low risk pregnancies: Prospective study in midwifery practices in the Netherlands. BMJ Volume 313, Issue 7068, 1996, Pages 1309-1313

Wiegers TA. (1998) *Home or hospital birth: a prospective study of midwifery care in the Netherlands.* European Journal of Obstetrics and Gynecology and Reproductive Biology 79 (1998) 139-141

Williams JG., Mann RY. (2002) *Hospital Episode Statistics: time for clinicians to get involved?* Clin Med 2002; 2:34–7.

Wong GY. and Mason WM. (1985) *The Hierarchical Logistic Regression Model for Multilevel Analysis.* J. Am. Statist. Assoc. 80:513-23.

Wright A., Ball J., Seccombe I., Stock J. (1993) *Mapping team midwifery: a report to the Department of Health.* Brighton: Institute of Manpower Studies, 1993

Workforce Review Team (2009) Workforce Review Team Annual Report 2007/8. Winchester: NHS Workforce Review Team.

Winterton Report (1992).

Yelland J., Rayner J., Lumley J., Davey MA. (2006) *Staffing in postnatal units: is it adequate for the provision of quality care? Staff perspectives from a state-wide review of postnatal care in Victoria, Australia*. BMC Health Serv Res. 2006 Jul 4; 6:83

Zander L. and Chamberlain G. (1999) *ABC of labour care place of birth.* British Medical Journal 318: 721-3.

Zhang WH., Alexander S., Bouvier-Colle MH., Macfarlane A. (2005) *Incidence of severe pre-eclampsia, postpartum haemorrhage and sepsis as a surrogate marker for severe maternal morbidity in a European population-based study: the MOMS-B survey.* BJOG 2005, 112(1):89-96

APPENDIX I

Text Box: National Service Framework (NSF) for Children, Young People and Maternity Services (2004)

Markers of Good Practice

1. All women are involved in planning their own care with information, advice and support from professionals, including choosing the place they would like to give birth and supported by appropriately qualified professionals who will attend them throughout their pregnancy and after birth.

2. Maternity services are proactive in engaging all women, particularly women from disadvantaged groups and communities, early in their pregnancy and maintaining contact before and after birth.

3. All services facilitate normal childbirth wherever possible, with medical interventions recommended only when they are of benefit to the woman and/or her baby.

4. Maternity services are commissioned within a context of managed care networks and include a range of provision for routine and specialist services for women and their families' e.g:

Routine ante-natal and post-natal care services;

Services for women with more complex pregnancies who may require multi disciplinary or multi-agency care;

Services for women who request support for coping with domestic violence;

Services for disabled women;

Services for women and their partners who request support to stop smoking;

Services for women and their partners who are substance misusers; and

Services for women and their partners who have mental health problems.

5. Managed maternity and neonatal care networks include effective arrangements for managing the prompt transfer and treatment of women and their babies experiencing problems or complications.

6. All women and their babies receive treatment from health care professionals competent in resuscitation for both mother and infant, newborn examination and in providing breastfeeding support. Services promote breastfeeding, whilst supporting all women whatever their chosen method of feeding.

7. Women who use local maternity services are involved in improving the delivery of these services, and in planning and reviewing all local hospital and community maternity services.

APPENDIX II

Midwifery staffing in varied birth settings based on case mix categories to provide the standard of one-to-one care in labour (adapted from Ball 2006 and RCOG 2007, p.29)

Setting	Birthrate Plus Case Mix Category	Definition of category	Midwife-to-woman standard ratio	MCA to midwife ratio
Home	I & II	Low risk: midwifery care; 37–42 weeks of gestation, normal birth, no ntervention, no epidural, good birth weight and Apgar	1 WTE midwife to 1 woman	1 MCA for team of 6 midwives
Birth centre	I & II	Low risk: midwifery care; 37–42 weeks of gestation, normal birth, no intervention, no epidural, good birth weight and Apgar	1 WTE midwife to 1 woman	1 MCA for team of 6 midwives
Obstetric unit based on case mix categories, not dependent on size or setting	I & II	Low risk: midwifery care; 37–42 weeks of gestation, normal birth, no intervention, no epidural, good birth weight and Apgar	1 WTE midwife to 1 woman	1 MCA for 6 midwives each shift to cover diverse duties (non- midwifery)

III	Moderate degree of intervention: induction, fetal monitoring, instrumental birth, third degree tear, preterm birth	1.2 WTE midwives to 1 woman	1 MCA for 4 midwives each shift to cover diverse duties (non- midwifery)
IV	Higher risk/higher choice or need; normal birth with epidural for pain relief, elective caesarean sections, post-birth complications	1.3 WTE midwives to 1 woman	1 MCA for 4 midwives each shift to cover diverse duties (non- midwifery)
V	Highest risk including emergencies; emergency caesarean sections, medical or obstetric complications, multiple births, stillbirths, severe pregnancy-induced hypertension	1.4 WTE midwives to 1 woman	1 MCA for 4 midwives each shift to cover diverse duties (non- midwifery)

APPENDIX III

The following medical and non-medical staff groups were requested from the NHS IC as part of the SDO project. The request was for both headcount and FTE by Agenda for Change pay band (or equivalent), O&G grade (medical staff), age, gender, ethnicity and nature of contract for each NHS Trust and SHA for the period 2000-2011.

1) Non-medical staff in Maternity

Non-medical staff in Maternity

NAC	Nurse Consultant Maternity services
NCC	Modern matron Maternity services
N0C	Manager Maternity services
N1C	Registered nurse - Children Maternity services
N2C	Registered midwife Maternity services
N2J	Registered midwife Education staff
N3H	Health visitor Community services
N6C	Other 1st level Maternity services
N7C	Other 2nd level Maternity services
N8C	Nursery nurse Maternity services
N9C	Nursing assistant/ auxiliary Maternity services
H1C	HCA Maternity
H2C	Support worker Maternity

2) Medical Staff by Grade in Obstetrics and Gynaecology

Grades of Medical Staff

Consultant
Associate Specialist
Specialty Doctor 1
Staff Grade
Registrar Group
Senior House Officer
Foundation Year 2
House Officer and Foundation Year 1
Hospital Practitioner

Clinical Assistant
Other Staff

THE FOLLOWING NON-MEDICAL STAFF GROUP DEFINITIONS ARE TAKEN FROM THE NHS IC OCCUPATIONAL CODE MANUAL:

http://www.ic.nhs.uk/article/2268/NHS-Occupation-Codes

http://www.ic.nhs.uk/CHttpHandler.ashx?id=11194&p=0

- **Nurse Consultants:** Those staff who are specifically appointed as such. They are experts within their area of clinical practice; provide professional leadership; work towards research and provide a function for education and professional development within their specialist clinical area. They work (unlike managers) directly with patients, clients or communities for at least 50% of their time. Typically Agenda for Change Band 8a and above.
- **Modern Matron:** A senior nurse with high level competencies and extensive clinical, leadership and management experience who has clinical and professional authority and responsibility for standards of professional practice and patient services in one or more service delivery areas. Typically Agenda for Change Band 8a.
- **Manager:** A nurse, midwife, or health visitor who has overall responsibility for budgets, manpower or assets or who is held accountable for a significant area of work and who has little or no direct clinical involvement; the post occupied would require the person to hold a statutory nursing, midwifery or health visiting qualification. Typically Agenda for Change Band 8 and above.
- **Registered Midwife:** An employee who holds a qualification as a Registered Midwife and who occupies a post where such a qualification is a requirement. Not below Agenda for Change Band 5 and usually Band 6 or above.

- **Other 1st Level (Level 1 - Sub Part 1):** First level nurses are registered nurses who hold a current and valid registration with the NMC under Level 1 Nurses Sub-part 1 of the register, but do not fulfil the criteria outlined in notes 1 to 10. Code according to their general area of work. Not below Agenda for Change Band 5.
- **Other 2nd Level (Level 2 - Sub Part 2):** Also referred to as 'enrolled' nurses, 2nd level nurse training is no longer provided. They are registered with the NMC under Level 2 Nurses Sub-part 2 of the register. There are likely to be few 2nd Level nurses within an organisation, as most nurses now hold a 1st level qualification. Second Level nurses can undertake a conversion course to become 1st Level nurses. More information about nurse registration and conversion courses is available on the NHS Careers website.
- **Children's Nurse:** An employee who holds the Registered Children's Nursing Certificate under Level 1 Nurses Sub-part 1 of the Nursing and Midwifery Council Register (NMC) and who occupies a post where such a qualification is a requirement. Not below Agenda for Change Band 5.
- **Nursing Assistant/Auxiliary:** Employees who are not required to hold any of the qualifications specified above who form part of the nursing workforce. This group also includes Cadet Nurses, who would not be accurately recorded as Pre-Registration Learners.
- **Nursery Nurse:** An employee who is employed as a nursery nurse and hold, or working towards, a relevant child care qualification including CACHE Level 3 Diploma in Child Care and Education, BTEC National Diploma in Children's Care, Learning and Development, NVQ Level 3 in Children's Care, Learning and Development. Information on relevant qualifications can be found on the Children's Workforce Development Council (CWDC) qualification list. (https://secure.cwdcouncil.org.uk/eypqd/qualification-search)
- **Health Care Assistants (HCA):** Code as **HCA** those staff who are trained, or under training in the various competencies related to their job. This training

might be through NVQ or other local HCA training. Support staff without formal NVQ or local HCA training should be coded as H2*.

THE FOLLOWING MEDICAL STAFF GROUP DEFINITIONS ARE TAKEN FROM THE NHS IC, MARCH 2011, P.14

https://catalogue.ic.nhs.uk/publications/workforce/numbers/nhs-staf-non-medi-2000-2010/nhs-staf-non-medi-cens-bull-2000-2010-rep.pdf

- **Career grades -** The component grades of this group are consultant, specialty doctor, associate specialist and staff grade.
- **Doctors in training and equivalents -** The component grades of this group are registrar group, senior house officer, specialty registrars (StRs) who are on fixed term specialty training appointments (FTSTAs), house officers, foundation programme doctors years 1 and 2 and other staff working at equivalent grades that are not in an educationally approved post.
- **Registrar group -** The component grades of this group are specialist registrars (SpRs), senior registrars, registrars, specialty registrars (StRs) who are on run through specialist training (ST grades) and other staff working at equivalent grades that are not in an educationally approved post.

APPENDIX IV

Box 1. Suggested reporting parameters for studies using HES data (from Sinha et al, 2012:5, Table 4)

1. Clearly identify the study as using retrospective administrative data in the title or abstract (a).
2. Use the term 'Hospital Episode Statistics' in the abstract and avoid use of alternative eponyms such as 'Health Episode Statistics' (b).
3. Clearly distinguish between HES (or equivalent centralized data) and PAS (or equivalent locally collected) data in the methods section of the paper (a,b).
4. Provide a rationale for selecting the study period and state that this was decided before data extraction (a).
5. Explicitly describe the parameters used in selecting participants and exposures, for example (a):
   - Listing all ICD and OPCS (or other) codes used.
   - Any inclusion or exclusion criteria and the order in which they were applied.
6. Clearly describe any steps taken in cleaning the data set, for example (a):
   - Identification of missing or invalid data.
   - Identification of duplicate records.
   - How missing, invalid or duplicate records were handled in analyses?
7. For studies involving more than 1 year of data, provide a statement that year to year data quality variations were assessed (a).
8. Explicitly report use of the HES_ID (or equivalent identifier) to create continuous in-patient spells from composite episodes and separate provider spells (a,b).
9. Clearly distinguish between whether episodes, spells, patients or procedures were counted and how these were defined (a).
10. For outcomes such as mortality, link HES records to other databases such as the Office of National Statistics (or equivalent national mortality register) to improve consistency (a,b).
11. For outcomes such as readmissions and re-interventions, take account of denominator changes with time (e.g. survival analysis) or provide a statement of the implications if this was not done (a).
12. For studies comparing health-care providers, provide some evidence of external case validation or a statement of the implications if this was not done (a).
13. In statistical modelling, clearly state the rationale for choice of candidate risk factors from available case-mix variables (a).
14. In statistical modelling, screen for evidence of non-additive risk relationships between case-mix variables (a).
15. Clearly state the presence or absence of any conflicts of interest and any sources of funding (a).

(a): Generic aspects with applicability to data sets from other countries.
(b): Specific aspects with applicability to HES data.

Box 2: A list of core issues related to the quality and coverage of HES maternity data.

The following list is taken from both Box 2, in Murray et al. (2012:5) and from HSCIC (2011:9; NHS maternity statistics 2010/11 explanatory notes): https://catalogue.ic.nhs.uk/publications/hospital/maternity/nhs-mater-eng-2010-2011/nhs-mate-eng-2010-2011-rep.pdf

- Stand-alone maternity systems in around 20 hospitals are not linked to their patient administration system, from which HES data are obtained (via the Secondary Uses Service);
- Some hospitals return data on the number of birth episodes or delivery episodes but not both;
- Transfer of maternity information between systems leaves scope for data errors and shortfalls;
- Stillbirths are not reliably recorded in every hospital and are not allocated an NHS number;
- Lack of a priori definitions for data variables resulting in inconsistencies in data entry;
- Use of aggregate or categorized fields rather than the raw data;
- Trusts identifying a high number of maternity beds available, but not recording any information about deliveries or births;
- Trusts identifying that they have no maternity beds available, but recording a high number of birth and delivery episodes;
- Some trusts have space in their maternity system to record nine birth tails, whereas other systems have space for 18. As deliveries, miscarriages and abortions are all recorded in the birth tail, there are cases where nine tails is not enough to record all of the relevant data.

APPENDIX V

NICE Clinical Guideline 55 Intrapartum care - Factors affecting place of birth

Table 1a Medical conditions indicating increased risk suggesting planned birth at an obstetric unit

Disease area	**Medical condition**
Cardiovascular	Confirmed cardiac disease Hypertensive disorders
Respiratory	Asthma requiring an increase in treatment or hospital treatment Cystic fibrosis
Haematological	Haemoglobinopathies – sickle-cell disease, beta-thalassaemia major History of thromboembolic disorders Immune thrombocytopenia purpura or other platelet disorder or platelet count below 100 000 Von Willebrand's disease Bleeding disorder in the woman or unborn baby Atypical antibodies which carry a risk of haemolytic disease of the newborn
Infective	Risk factors associated with group B streptococcus whereby antibiotics in labour would be recommended Hepatitis B/C with abnormal liver function tests Carrier of/infected with HIV Toxoplasmosis – women receiving treatment

	Current active infection of chicken pox/rubella/genital herpes in the woman or baby Tuberculosis under treatment
Immune	Systemic lupus erythematosus Scleroderma
Endocrine	Hyperthyroidism Diabetes
Renal	Abnormal renal function Renal disease requiring supervision by a renal specialist
Neurological	Epilepsy Myasthenia gravis Previous cerebrovascular accident
Gastrointestinal	Liver disease associated with current abnormal liver function tests
Psychiatric	Psychiatric disorder requiring current inpatient care

Source: NICE Clinical Guideline 55 Intrapartum care Table 3.7 page 64, reproduced with permission.

Table 1 b Other factors indicating increased risk suggesting planned birth at an obstetric unit

Factor	**Additional Information**
Previous complications	Unexplained stillbirth/neonatal death or previous death related to intrapartum difficulty Previous baby with neonatal encephalopathy Pre-eclampsia requiring preterm birth Placental abruption with adverse outcome Eclampsia Uterine rupture Primary postpartum haemorrhage requiring additional treatment or blood transfusion Retained placenta requiring manual removal in theatre Caesarean section Shoulder dystocia
Current pregnancy	Multiple birth Placenta praevia Pre-eclampsia or pregnancy-induced hypertension Preterm labour or preterm prelabour rupture of membranes Placental abruption Anaemia – haemoglobin less than 8.5 g/dl at onset of labour

	Confirmed intrauterine death Induction of labour Substance misuse Alcohol dependency requiring assessment or treatment Onset of gestational diabetes Malpresentation – breech or transverse lie Body mass index at booking of greater than 35 kg/m² Recurrent antepartum haemorrhage
Fetal indications	Small for gestational age in this pregnancy (less than fifth centile or reduced growth velocity on ultrasound) Abnormal fetal heart rate (FHR)/Doppler studies Ultrasound diagnosis of oligo-/polyhydramnios
Previous gynaecological history	Myomectomy Hysterotomy

Source: NICE Clinical Guideline 55[4] Intrapartum care Table 3.8 page 65, reproduced with permission.

Table 1 c Medical conditions indicating individual assessment when planning place of birth

Disease area	Medical condition
Cardiovascular	Cardiac disease without intrapartum implications
Haematological	Atypical antibodies not putting the baby at risk of haemolytic disease Sickle-cell trait Thalassaemia trait Anaemia – haemoglobin 8.5–10.5 g/dl at onset of labour
Infective	Hepatitis B/C with normal liver function tests
Immune	Immune Non-specific connective tissue disorders
Endocrine	Unstable hypothyroidism such that a change in treatment is required
Skeletal/neurological	Spinal abnormalities Previous fractured pelvis Neurological deficits
Gastrointestinal	Liver disease without current abnormal liver function Crohn's disease Ulcerative colitis

Source: NICE Clinical Guideline 55[4] Intrapartum Care Table 3.9 page 65, reproduced with permission

Table 1d Other factors indicating individual assessment when planning place of birth

Factor	Additional Information
Previous complications	Stillbirth/neonatal death with a known non-recurrent cause Pre-eclampsia developing at term Placental abruption with good outcome History of a previous baby more than 4.5 kg Extensive vaginal, cervical or third or fourth degree perineal trauma Previous term baby with jaundice requiring exchange transfusion
Current pregnancy	Antepartum bleeding of unknown origin (single episode after 24 weeks of gestation) Body mass index at booking of 30-34 kg/m^2 Blood pressure of 140 mmHg systolic or 90 mmHg diastolic on two occasions Clinical or ultrasound suspicion of macrosomia Para 6 or more Recreational drug use Under current outpatient psychiatric care Age over 40 at booking
Fetal indications	Fetal abnormality
Previous gynaecological history	Major gynaecological surgery Cone biopsy or large loop excision of transformation zone Fibroids

Source: NICE Clinical Guideline 55[4] Intrapartum Care Table 3.10 page 66, reproduced with permission

APPENDIX VI

VARIABLES IN THE MODELS

A full list of all the variables used in the models, with details of their coding, are presented below

Variable Names, Coding and Description

Variable name	Description and codes
Emergency CS	Emergency CS (1=yes; 0=no); response variable
Instrumental Delivery	Instrumental Delivery (1=yes; 0=no); response variable
Normal Birth	Normal Birth (1=yes; 0=no); response variable
Delivery ID	Woman ID (Level-1 unit)
Trust ID	NHS Trust ID (Level-2 unit)
Maternal Age (grouped in years)	Age of mother at time of baby's birth (1=15-19 (Ref); 2=20-24; 3=25-29; 4=30-34; 5=35-39; 6=40-44); 7=Missing
IMD	Index of Multiple Deprivation in quintiles (1=Least deprived (Ref); 5=Most deprived); 6=Missing
Ethnicity	Ethnicity of mother aggregated (1=White (Ref); 2=Asian; 3=Black Afro-Caribbean; 4=Chinese, Mixed, Other; 5= Unknown)
NICE Any Risk	NICE (2007) definition of risk plus individual assessment (1=High Risk; 0=Low Risk), reversed in normal birth model
GA	Gestational Age in weeks (GA between 37 weeks and 41 weeks (Ref)=1; GA >=42 weeks=2; 3=Missing
BW	Birth Weight in grams (BW 2500g-4500g (Ref)=1; BW<2500g=2; BW>4500g=3); 4=Missing
London Trust	1=London trust; 0=Not London Trust (Ref) (Level-2 variable)
FT Status (Sept 2010)	1=Foundation trust; 0=Not a Foundation trust (Ref) (Level-2 variable)
University Hospital	1=Teaching trust; 0=Not a Teaching Trust (Ref) (Level-2 variable)
Trust OU/AMU/FMU	1=Trusts with Obstetric Unit (OU) only (Ref); 2=Trusts with OU and AMU (Alongside Midwifery Unit); 3=Trust with OU and FMU (Freestanding Midwifery Unit); 4=Trust with OU and AMU and FMU (Level-2 variable)
ZMidwives FTE/birth	Standardized ratio of Midwives FTE/birth (Level-2 variable)
ZConsultants FTE/birth	Standardized ratio of Consultants O&G FTE/birth (Level-2 variable)
ZDoctors FTE/birth	Standardized ratio of Doctors FTE/birth (Level-2 variable)
ZHCA FTE/birth	Standardized ration of Healthcare Assistants FTE/birth (Level-2 variable)

APPENDIX VII

LITERATURE REVIEW STUDIES

Study	Country	Study design	Study population	Outcomes assessed	Key participants inclusions/exclusions
STRUCTURE/OUTCOMES STUDIES					
Maternity staffing studies					
Joyce et al. 2002	UK (England)	Cross-sectional	540,834 births	CS, epidural use in labour, IVD	none
Hall et al. 2009	US (Utah)	Retrospective, descriptive analysis	900 women who delivered in 3 hospitals	Duration of three stages of labour; labour complications; CS; fetal distress; patient length of stay and cost of care	none
Ashcroft et al. 2003	UK (England)	Prospective semi-structured observational	All midwives on labour wards of 7 maternity units	'Latent failure' in relation to midwifery staffing, deployment and training	none
Gerova et al. 2010	UK (England)	Cross-sectional	615,042 women	Maternal readmission within 28 days	none
Place of birth studies					

Hodnett et al. 2010	Denmark, Australia, Ireland, UK, Canada, Scotland, Sweden	Cochrane review of 9 randomized or quasi-randomized controlled trials which compared the effects of an alternative institutional birth environment to conventional maternity ward care	10,684 women	Spontaneous vaginal births, IVD, CS, PPH, maternal satisfaction, epidural analgesia, oxytocin augmentation of labour, episiotomy, intrapartum analgesia/anaesthesia, perinatal and maternal severe morbidity and mortality, breastfeeding at six to eight weeks, five-minute Apgar score less than seven, admission to neonatal intensive care unit	
Olsen 1997	USA, Australia, England, Netherlands, Australia,	Meta-analysis based on 6 observational studies	24,092 women	Safety of planned home births, induction, augmentation, IVD, CS episiotomy, and perinatal outcomes (five-minute Apgar score), perinatal mortality	Low risk women

	Switzerland				
Wiegers 1996	Netherlands	Prospective study	1836 women, 97 midwives	Planned home/hospital births - Perinatal outcome index; obstetric outcomes – PPH, perineal tears, length of labour,	Low risk women
de Jonge et al. 2009	Netherlands	Retrospective study using routine data	500,000 women	Perinatal mortality and morbidity	Low risk women
Janssen et al. 2002	Canada	Prospective cohort study	2176 women	Apgar scores, neonatal mortality, meconium aspiration; epidural analgesia, induction, augmentation, episiotomy	
Symon et al. 2009	Scotland	Anonymised matched cohort study	8676 women	Normal birth, spontaneous onset of labour, use of analgesia, perineal trauma, breastfeeding; prematurity, low birth weight, admission to neonatal intensive care	
Birthplace 2011	England	Prospective cohort study	60,000 women	Normal birth, augmentation, ventouse or forceps delivery, intrapartum caesarean section, episiotomy, adverse maternal outcomes - 3rd/4th	Low risk women

				degree perineal trauma, blood transfusion and admission to higher level care. Safety, models of care and cost-effectiveness of different settings for birth	
PROCESS/OUTCOMES STUDIES					
Continuity of care, one-to-one midwifery care, models of care					
Page et al. 1999	England	Prospective comparative study	902 women	epidural, episiotomy, perineal tears, women's satisfaction, normal birth, cost	
Biro et al. 2000	Australia	RCT	1000 women	augmentation, electronic fetal monitoring, epidurals and episiotomy, length of stay, perinatal mortality	
Hodnett et al. 2013	16 countries	Cochrane review of 22 RCT	15,288 women	duration of labour, epidural analgesia or anaesthesia, spontaneous vaginal birth, instrumental delivery, CS, women experience	
Sandall et al. 2013	Australia, Canada, UK, New Zealand	Cochrane review of 13 RCT	16,242 women	analgesia; episiotomy, instrumental births; spontaneous vaginal births; CS, preterm birth, breastfeeding.	